(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000) 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

09/496.914 (CIP) US 3 February 2000 (03.02.2000) Filed on US 09/560,875 (CIP) Filed on 27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston,

MA 02111 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,

TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

5

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

5

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

5

10

15

20

25

30

35

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, butilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

10

15

20

25

30

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

5

10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

5

10

15

20

25

30

35

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4²⁰ possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match $(1 \div 4^{25})$ times the increased probability for mismatch at each nucleotide position (3×25) . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

10

15

20

25

30

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

5

10

15

20

25

30

35

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (*i.e.*, hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (*i.e.*, washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

15

5

10

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (*i.e.* specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

5

10

15

20

25

30

35

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

10

15

20

25

30

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

10

15

20

25

30

35

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

25 4.3 ANTISENSE

10

15

20

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

10

15

20

25

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

. 15

20

25

30

35

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556;

Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition,

oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

10

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., Basic Methods in Molecular Biology (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

10

15

20

25

30

35

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5

10

15

20

25

30

35

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBatTM kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearlTM or Cibacrom blue 3GA SepharoseTM; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

20

25

30

35

10

5

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

10

15

20

25

30

35

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e,g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

5

10

15

25

30

35

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

10

15

. 20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

10

15

20

25

30

35

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

5

10

35

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation,

Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells 20 include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 25 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology, J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 30 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

5

10

20

25

30

35

A

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation. manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5

10

15

20

25

30

35

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis 10 and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with 15 irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or 20 treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and 25 paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) . as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

20

25

30

10

5

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

10

15

20

25

30

35

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No.

15 WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

5

10

15

20

25

. 30

35

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus. rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

10

15

20

25

30

35

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by

removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed

APCs either expressing a peptide of the present invention or together with a stimulatory form of
a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the
patient. Another method of enhancing anti-viral immune responses would be to isolate infected
cells from a patient, transfect them with a nucleic acid encoding a protein of the present
invention as described herein such that the cells express all or a portion of the protein on their
surface, and reintroduce the transfected cells into the patient. The infected cells would now be
capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β₂ microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

15

20

25

30

35

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

5

10

15

20

25

30

35

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polypucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

ý

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

10

15

20

25

30

35

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

30

35

5

10

15

20

25

4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

5

10

15

20

25

30

35

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

30

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

5

10

15

20

25

30

35

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- 10 (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

5

15

20

25

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particularneurotoxins; and
 - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

30

35

- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
 - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without

5 limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of

10 hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

15

20

25

30

35

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

10 4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

15

20

25

30

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

5

10

15

30

35

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). 20 Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or 25 amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

5

10

15

20

25

30

35

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

35

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules. liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity.

Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

10

15

20

25

30

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

5

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials 5 are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above 10 mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl 15 cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, 20 hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on 25 total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other 30 agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

35

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

15

5

10

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC₅₀ as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD₅₀ and ED₅₀. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

5

10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG₁, IgG₂, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A. synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

25

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

5

35

10 The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins. immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigen-15 binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the 20 corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable 25 domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 30 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al., (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 Fab Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab)2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab)2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

5.13.5 Bispecific Antibodies

5

10

15

20

25

30

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., 1991 EMBO J., 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., <u>Science</u> 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

10

15

20

25

30

35

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

5

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention.

Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HTV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

5

10

15

20

25

30.

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

10

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that annual to a polynucleotide of the invention under such conditions, and amplifying annualed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

4.18 SCREENING ASSAYS

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
 - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

10

15

20

25

30

35

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

10 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

10

15

20

25

30

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata et al., 1985; Dahlen et al., 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller et al., 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm₇), is then added to a final concentration of 10 mM 1-MeIm₇. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm₇, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook et al. (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*II, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviJI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme ($CviJI^{**}$), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a $CviJI^{**}$ digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that $CviJI^{**}$ restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

5

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm² and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

5

10

15

20

25

30

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

5 Novel Contigs

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (*i.e.*, Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

5

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
			976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
			217 225 238 271 317 404 446 469 503
	j		513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
			1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
			1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
		,	147 188 197 208 225 227-239 250 300-
			303 312 316 328-331 340 357-362 374
		•	380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550
			566 571 577 585 590 594 598 634 641
i		}	658 666 683 725 742 764 767 786 801
		ł	805 810 823 826 829 831 836 841 887-
			923 927 934 943 950-951 963 976 995
			1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
			1128 1142 1162 1181-1192 1199 1204
		Ť	1218-1219 1225 1232 1253 1267 1271-
3 10 3		1,770	1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
	[566 596 663 670 746 798 816-819 876
14. 3		477001	892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes	Clarate	ADDOOD	740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
	·		240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
			1003 1067-1070 1118 1156 1193-1200
adult heart	GIBCO	AHR001	1325
adult licalt	GIBCO	Ankou	38 49 71-72 74-77 79 92 99 101 111 118 129 132 138 151 158-163 182 195-
			,
			203 215 217 238 264 269 353 384 398 408 434-439 446 504 512-513 519 537
•			562-573 577 611-614 616-619 658 661
			671-672 722 734 757-773 815 828-835
			874 891 898 919 926-927 976 988
		1	1021 1037 1041 1062 1067 1071 1080
		į.	1083 1093 1122 1131 1185 1201 1254
	İ		1308 1331 1335
adult kidney	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103-
acuit riuncy	GILCO	AKTOO!	107 111 119-120 138 151 157 215 217-
			218 238 250 264 294 304 384 404 440
			446 454 477 504-505 509 514 518-519
	-		535 537 564 574-583 620-627 639 653
		,	673-675 705 753 789 831 844 851 859
			877 909 918 927 956 963 976 1067
			1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	11-12 41 49 111-112 215-217 294 316
wan rimey	mainoRen	AA 1002	446 487 564 575 844 868 910 927 976
			1116
adult lung	GIBCO	ALG001	8 101 111 151 187 402 446 490 514
accent tung	I ODDCO	TATOMI	0 101 111 131 107 402 440 490 314

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			518 537 545 549 580 582 592 594 634
		1	640 651-652 676-678 725 851 873 918
		1	952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
тушри посо	Cioniaci	1221001	545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
Aomia uset	ODECO	ALVOOI	519 537 564 653 683-684 698 753 798
		1	813 833 840 858 927 976 1038-1039
4 4. 10		1	1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
		·	505 657 675 714 753 832 844 941-942
			976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
		1	104 111 120 122-125 138 140 143-149
		1	151 188-190 207-212 215-217 238 264
		· C	316 384 409 440 445-446 496 504 512
		1	514 518-519 535 537 549-550 564 566
		1	571 580 582 600 618 638 657 667 681
			685-697 699 705 722 735-744 761 771
		ŀ	815 833 842-865 868 875-876 918 926-
		Į.	927 950 952 963 976 1023 1042 1048
		ļ	1051 1059 1072 1076 1083 1117 1120
		ļ	1124 1131 1144 1174 1224 1268 1331
		ļ	1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
aduit spieen	GIBCO	ASPOUL	294 414 446 477 504 514 534 545 549
		1	,
		1	592 722 873 883 952 976 1041-1042
			1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215
		1	238 446 497 537 642 701-706 811 877
			927 962 976 1083 1117 1131
adult bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
			592 607 706 873 952 1178 1329-1335
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
			147 151 164-174 213-215 238 305-307
		1	374 404 446 460 466 516 519 534 538-
			541 544-546 549-554 566 584 586 592
			596 607 610 628-629 643-645 652 707-
	}		708 774-789 844 866-871 873 919 927
		1	952 963 976 998 1034 1042 1064 1083
	•		1085 1120 1132 1152 1225 1229 1268
		1	1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
	0.00.00		210 317 510-511 545 549 581 598 628
	1	J	638 724 766 789 844 860 868 873 919
		•	927 952 963 968 976 1042 1111 1141
			1160-1161 1229 1266 1346
hono morno	Clontech	BMD004	111 238 282 549 1083
bone marrow			
adult colon	Invitrogen	CLN001	52 260 264 299 494 536 545 564 592
		j	844 873 877 952 976 1042 1152 1268
		ļ	1336-1337
adult cervix	BioChain	CVX001	49 51 129 132 151 205 207 238 332-
			335 365-367 392-401 440 466 470-471
			518 537 597 629 832 877 927 976 1006
			1085 1117 1129-1134 1192 1202-1205
		<u> </u>	1219 1309-1328
diaphragm	BioChain	DIA002	74 976 1083
			

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
			138 151 204-206 215-217 238 269 316
ļ	}		414 433 505 510 513 550 555 580 582
	i		596 675 722 745 798 814 836-841 851
Ì		-	918 976 1041 1043 1073 1083 1131
			1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic		1
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		}
of chromosome 8 Genomic clones	Research	T272 (00 (205 005
from the short arm	Genomic DNA from Genetic	EPM004	525 927
of chromosome 8	Research	j	
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic	ELMIOO) 331
of chromosome 8	Research		
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
	0.02.70	1 21.000	225 271 317 319 336 359 368 405-414
1			519 550 571 594 686 715 722 764 824
		}	829 836 859 909 927 943 947 963 1057
			1067-1068 1104 1135-1140 1162 1206-
<u>}</u>	1		1207 1235 1268 1288 1307-1308 1319
	Ì		1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
			535 683 761 798 820-827 844 876 909
<u> </u>	 		963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537 550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD002	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
	2	- 2000	859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
,	University		69-71 74 77 79 87-90 101 107 110-111
			114 120 128-131 138 140 147 150-155
			197 210 215 217 225 238 312 367 384
		·	414 440 446 460 468 483 496 504-507
			511-515 518-519 523 533-535 537 541
	1		544-545 547-550 555-560 564 566 571
			577 582 585-586 598 636 646-647 649
		1	652 664 698 709-710 714 722-723 731
	}	,	735-736 746-753 761 784 798 823 829
		•	832 844 851 858-859 868 873 876 898
		1	927 943 949 952 963 976 984 1002
			1021 1023 1040 1042 1044 1050 1083
			1093 1116 1120 1129 1131 1144 1174 1217 1251 1254 1256 1302 1308 1311
			1217 1251 1254 1256 1302 1308 1311
fetal liver-spleen	Columbia	FLS002	8 36-37 41-46 49 54 64 71 74 79 101
romi ii.oiahioni	University	A LAUGOZ	111 120 129 147 207 210 215-216 238
	3		250 330 353 359 366 383-384 414 478
		·	505 508-509 511 515-524 534-535 537
			544-545 564 566 571 577 591 598 638
	<u> </u>		0.1.010007000011011071070000

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			663 671 698 714 722 725 727 751 798
ļ			851 859 873 876 909 927 949 952 983-
			984 1002 1023 1042-1044 1085 1095
	1	1	1131 1144 1178 1199 1233 1240-1270
<u> </u>	 		1331 1340
fetal liver-spleen	Columbia University	FLS003	64 535 976 1256
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
	and the open	1.2,001	580 722 730 749 844 918 943 976 1051
			1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
			425 535 537 577 598 614 836 857 1141
	,		1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
	1		225 264 316 405 422-429 488-494 496
		[519 534-535 537 566 675 732 859 876-
			877 898 947 949-950 963 976 1001
			1062 1076 1083 1117 1144 1165 1268
			1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
	1	1	316 446 495-503 519 521 534-535 537
	ł		582 634 691 877 883 927 944-950 963
	ĺ		976 1001 1075 1142-1143 1171 1218
fetal brain	CIDGO	VIEDO01	1243 1308
iciai brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
			138 145 151 188 197 207 215 238 264 271 294 316 367 414 440 446 466 504
·	ĺ		513-514 535 542-543 550 564 571 596
			635 648-654 675 711-715 722-723 798
	1		832 872 876 883 927 976 1095 1144
	l		1168 1171 1178 1211 1335
macrophage	Invitrogen	HMP001	238
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
	University		151 161 175-179 185 216-217 264 295
	1		299 308-310 371-373 462 476 504 511-
			513 533 537 564 566 571 655-657 662
			683 716-720 723 752 790-803 829 832
	•		858-859 876 898 909 949 976 1045-
			1047 1076-1087 1090 1093 1116 1122
			1144 1209-1213 1225 1233 1256 1319
		<u> </u>	1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
	University		519 566 655 714 794 918 943 976 1067
	 	TT) (000	1092-1093 1233
infant brain	Columbia	IBM002	311 472-473 753 1214
infant brain	University Columbia	IBS001	51 111 376 474 790 876 949 1144 1204
mant bram	University	103001	
lung, fibroblast	Strategene	LFB001	1221 151 316 462 514 534 582 675 939 1131
lung tumor	Invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156
control	m m ugui	201002	215 217 269 280 296 337 374-375 384
	}	j į	404 446 454 475-480 498 514 518-519
	I	<u> </u>	522 537 545 564 577 597 653 658 705
	<u> </u>		· - •
			721-724 754-756 779 859 868 872-874
			· - •

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
		<u> </u>	1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
		<u> </u>	634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
			147 151 212 215 218 238 252 288 312-
Ì			314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
	1	}	564 566 571 577 580 582 587-609 615
			632-638 658-659 698 714 725-728 832
	1		836 841 859 866 873-874 882-883 918-
	}		919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
ļ. ,	-	1110000	1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
	1	1	657 660 729 873 883 927 952 963 1008
) (-1	0	2001.004	1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL 1424	}		919 929 939 952 976 1071 1118 1218
	Touris	1 0 0001	1235 1245
mammary gland	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
		Ì	217 250-256 264 297-299 305 377-378
	ļ	,	398 446 481-486 505 512 537 545 549
			571 592 725 730-733 816 829 836 844 868 873 876-877 898 926 943 951-960
			963 976 995 1034 1042 1048 1054-
		j	1055 1076 1083 1091 1093 1116-1117
	[1124 1152 1302
induced neuron cells	Strategene	NTD001	39 101 111 138 238 361 1225 1251
monceo nemon cens	Sualegene	NIDOOI	1319
retinoid acid induced	Strategene	NTR001	74 225 976
neuronal cells	Strategene	NIKOUI	14 223 976
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
prostate	Cionicai	rk1001	1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
	1		1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
			976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
	ŀ].	545 592 660 789 836 866 873 927 952
} .		}	963 967-978 1042 1120 1152 1223-
			1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
			270 343-344 353 379 516 537 566 740
		1	828 927 976 979-994 1092 1153-1159
			1225 1250
adult spleen	Clontech	SPLc01	698 859 1042
stomach	Clontech	STO001	210 238 271-272 537 580 705 918 952
			995 1171
thalamus	Clontech	THA002	61 219-220 273-276 312 315 330 596
	L	<u> </u>	963 996-1007 1059 1093 1160-1162
thymus	Clonetech	THM001	8 120 151 208 221 316-317 353 639
1			750 867 874 878-881 927 963 1023
			1083 1094-1096 1124
thymus	Clontech	THMc02	8 61 114 129 132 210 225 231 306
			317-319 336 340 359 380 398 446 448-
		}	463 512 519 545 554 587 598 698 724-
			725 789 812 836 868 873 927 947 952

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
			1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
			210 217 222 253 264 271 277-286 294
			320-326 345-352 361 381-382 446 467
			483 514 534 549-550 564 578 602 649
	1		844 882-883 927 950 956 976 1008-
	ļ		1028 1076 1083 1117-1120 1142 1163-
		<u> </u>	1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
		1	545 592 611 873 883-884 927
			952 1029-1031 1042 1151-1152
			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
	1	Ì	885-886 976 1001 1032-1033
		1	1232

TABLE 2

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AP036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEO	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:	1	ì		Score	1
29	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	83	42
30	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
31	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	96	67
32	G03224	Homo sapiens	Human secreted protein, SEQ ID NO: 7305.	58	32
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.	2457	98
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID NO:110.	348	95
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein sequence SEQ ID NO:150.	982	90
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))	687	99
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Homo sapiens	CGI-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
43	X61048	Hydra sp.	mini-collagen	105	35
44	M76546	Helianthus annuus	hydroxyproline-rich protein	110	31
45	U82288	Caenorhabditi s elegans	Rac-like GTPase	139	70
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	118	58
47	AF090942	Homo sapiens	PRO0657	113	63
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	90	59
49	AJ005560	Mus musculus	SPR2B protein	72	56
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
57	AL031600	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
58	AF011417	Mus musculus	putative pheromone receptor	143	55
59	AF167320	Mus musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapiens	interferon regultory factor 7	263	96
61	X07984	Mus musculus	protein-tyrosine kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	791	98
63 64	U35376 AF265555	Homo sapiens Homo sapiens	repressor transcriptional factor ubiquitin-conjugating BIR-domain enzyme	485 785	65 74
	1		APOLLON		ļ <u>.</u> _
65 66	G03883 AF177390	Homo sapiens Manduca	Human secreted protein, SEQ ID NO: 7964. antennal specific membrane protein AMP	88 274	95 54
<u> </u>	1.00/2222	sexta	CONTROL		ļ.,,,
67 68	AB040800 AF030027	Homo sapiens Equine	SREB2 24	213	26
60	C02055	herpesvirus 4	Human geometral protein CEO ID NO. 7046	261	05
69	G02965 W75770	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
70 71	AB011135	Homo sapiens	Human oxidoreductase YTFO3. KIAA0563 protein	239	98 76
72	AB011133 AB014885	Homo sapiens Halocynthia roretzi	HrPOPK-1	813	78
	AF045454	Cavia	phospholipase B	955	73
73	11 015154	porcellus			

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1		Waterman	Identity
NO:	<u> </u>	<u> </u>		Score	1
		musculus			
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
		norvegicus			
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein	351	54
	1	<u> </u>	complex component TRAP240		
77	Y38422	Homo sapiens	Human secreted protein.	468	76
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha- l-I (hCavT3).	1357	99
79	Y14591	Human	APM-1 protein	767	100
13	1 14391	papillomaviru	Arwi-i protein	/6/	100 .
	ĺ	s type 68	·	{	1
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis	protein arginine N-methyltransferase-like protein	359	65
0.	74 000505	thaliana	protein argume in-incury idansie ass-like protein	337	1 05
82	L46815	Mus	DNA binding protein Rc	895	75
-	270012	musculus	2101 omong proton to	""	1 "
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	538	71
	}		designated HSCOP-6.		1
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272_7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid	156	48
		1	sequence SEQ ID NO:100.	1 .	
88	AJ225124	Mus	hyperpolarization-activated cation channel,	487	95
		musculus	HAC3	,	1
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93	AF170723	Homo sapiens	protein kinase STK10	401	53
94 .	X13292	Trypanosoma	GPI-phospholipase C (AA 1 - 358)	151	37
		brucei	 		ļ
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus	sodium channel protein I (aa 1-2009)	1775	92
97	AF134213	norvegicus		1995	99
98	G00838	Homo sapiens Homo sapiens	ubiquitin-specific protease Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	675	48
"	AI-021933	norvegicus	kinase	0/3	** °
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence	160	60
101	11000,0,0	Tromo saprens	difference at residue 58	100	1 00
102	U22829	Mus	P2Y purinoceptor	264	42
		musculus	 	:	'-
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled	516	99
			receptor-B3.		
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
104	Y87342	Homo sapiens	Human signal peptide containing protein HSPP-	343	57
105	10/344				i
	L		119 SEQ ID NO:119.		L
106	AF169312	Homo sapiens	119 SEQ ID NO:119. hepatic angiopoietin-related protein	212	67
106 107	AF169312 AF116657	Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310	74	67 52
106	AF169312	Homo sapiens Homo sapiens Escherichia	hepatic angiopoietin-related protein		
106 107 108	AF169312 AF116657 AE000401	Homo sapiens Homo sapiens Escherichia coli	hepatic angiopoietin-related protein PRO1310 sialic acid transporter	74 587	52 96
106 107 108	AF169312 AF116657 AE000401 Y38395	Homo sapiens Homo sapiens Escherichia coli Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10.	74 587 693	52 96
106 107 108	AF169312 AF116657 AE000401	Homo sapiens Homo sapiens Escherichia coli	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone	74 587	52 96
106 107 108 109 110	AF169312 AF116657 AE000401 Y38395 Y78801	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence.	74 587 693 182	52 96 100 94
106 107 108 109 110	AF169312 AF116657 AE000401 Y38395 Y78801 Z25535	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153	74 587 693 182	52 96 100 94
106 107 108 109 110	AF169312 AF116657 AE000401 Y38395 Y78801	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein	74 587 693 182	52 96 100 94
106 107 108 109 110 111 112	AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84.	74 587 693 182 464 274	52 96 100 94 85 51
106 107 108 109 110 111 112	AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td	74 587 693 182 464 274	52 96 100 94 85 51
106 107 108 109 110 111 112	AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365 AC007956	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td unknown	74 587 693 182 464 274 301 520	52 96 100 94 85 51 71 75
106 107 108 109 110 111 112 113 114 115	AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365 AC007956 M83738	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td unknown protein-tyrosine phosphatase	74 587 693 182 464 274 301 520 251	52 96 100 94 85 51 71 75 92
106 107 108 109 110 111 112	AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365 AC007956	Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td unknown	74 587 693 182 464 274 301 520	52 96 100 94 85 51 71 75

SEQ	Accession	Species	Description	Smith-	1%
ID ID	No.	Species	Description	Waterman	Identity
	No.				laenuty
NO:	 	 		Score	
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rettus	phosphatidylinositol 3-kinase	627	93
	1	norvegicus		}	1
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory	1646	94
		1	subunit precursor, PDPr	1	
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC	373	68
	00,55,2	Lionio sepiens	3.1.3.48}	3/3	~~
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	00
					88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi	contains similarity to C2 domains	219	29
		s elegans	·		1
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit	693	90
	1		4		
126	AB021861	Mus	apoptosis signal-regulating kinase 2	153	65
	1	musculus	1		
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter	807	97
	1.200210	Tromo suprous	hCNT3	100,	1"
128	M90360	Homo sapiens	protein kinase	220	73
128	D32202			1	
		Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IFI16b	496	67
131	AF201734	Mus	testis specific serine kinase-3	800	87
	<u> </u>	musculus			
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73	1157	87
	1	monage supposes	clone HSOEL25.	****	"
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674 2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	5041	99
	<u> </u>	<u></u>	similarity to P49205 (PID:g1345860)		1
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis	DNA-damage inducible protein DDII-like	147	55
	1	thaliana			i
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus	AMPA receptor binding protein	623	93
	1.20,0115	norvegicus	1 1 Land 11 1000 proteins	1 023	1 -5
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus	transmembrane receptor UNC5H2	578	84
144	087300		transmemorane receptor ONC3712	3/8	04
	1	norvegicus			
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein	727	92
			M160 precursor	L	<u> </u>
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia	HrsA	818	90
		coli		1	
149	M83316	Escherichia	pppGpp phosphohydrolase	915	95
		coli	Property	1	
150	AL163279	Homo sapiens	homolog to cAMP response element binding and	1261	99
150	100105277	Tromo sapions		1201	133
151	A P170067	Name and and	beta transducin family proteins	040	100
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	392	61
			ligand (clone 3TW).	[<u> </u>
153	AF151859	Homo sapiens	CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type 1	489	81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
		Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
156	I G00857	,		I	74
156	G00857		zinc finger protein	1 357	. / "
156 157	AF159455	Mus	zinc finger protein	352	1 ' '
157	AF159455	Mus musculus			<u></u>
157 158	AF159455 L76191	Mus musculus Homo sapiens	interleukin-1 receptor-associated kinase	537	76
157	AF159455	Mus musculus	interleukin-1 receptor-associated kinase putative gene, ankirin like, possible dual		<u></u>
157 158 159	AF159455 L76191 AP001743	Mus musculus Homo sapiens Homo sapiens	interleukin-1 receptor-associated kinase putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	537 670	76 98
157 158	AF159455 L76191	Mus musculus Homo sapiens Homo sapiens Rattus	interleukin-1 receptor-associated kinase putative gene, ankirin like, possible dual	537	76
157 158 159	AF159455 L76191 AP001743	Mus musculus Homo sapiens Homo sapiens	interleukin-1 receptor-associated kinase putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	537 670	76 98

SEQ	Accession	Species	Description	Smith-	T%
ID	No.	opcacs	Description	Waterman	Identity
NO:	140.	ļ		Score	Identity
	722069	¥ * * * * * * * * * * * * * * * * * * *	\ \(\frac{1}{2} \)		100
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
164	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
165	AF160798	Rattus norvegicus	calcium transporter CaT1	700	96
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus musculus	NIK-related kinase	197	43
169	AF252293	Homo sapiens	PAR3	596	44
170	U59429	Cricetinae	diacylglycerol kinase eta	481	82
171	AF035268	gen. sp. Homo sapiens		386	42
			phosphatidylserine-specific phospholipase A1		
172	AF127085	Mus musculus	semaphorin cytoplasmic domain-associated protein 3B	507	82
173	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	653	99
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55 .
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196 4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in	710	99
		1	codon)		
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor	175	80
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein	301	98
			sequence SEQ ID NO:42.		
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria parva	casein kinase II alpha subunit	364	50
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	90
195	W95631	Homo sapiens	Homo sapiens secreted protein gene clone hi968 2.	382	49
196	AF255614	Rattus	scaffolding protein SLIPR	680	99
197	AC021640	norvegicus Arabidopsis	putative phosphatidate phosphohydrolase	300	41
160	1.000000	thaliana		 	ļ
198	AF073967	Mus musculus	olfactory receptor	316	43
	1	domesticus		<u> </u>	<u> </u>
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79
205	S81752	Homo sapiens	DPH2L=candidate tumor suppressor gene	375	100
	1 302132	Troute achienz	cemerceme minor subbressor Rene		1,100

SEQ	Accession	Species	Description	Smith-	%
\mathbf{D}	No.			Waterman	Identity
NO:	<u> </u>	<u> </u>		Score	<u> </u>
			{ovarian cancer critical region of deletion}		<u> </u>
206	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209_	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific	541	82
	<u> </u>	<u> </u>	protein, calphotin.		L
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	99
212	U81035	Rattus	ankyrin binding cell adhesion molecule	471	69
	<u> </u>	norvegicus	neurofascin	500	<u> </u>
213	AF154846	Homo sapiens	zinc finger protein	798	56
214	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	933	93
015	47.1/2202	musculus	deli della d	523	89
215 216	AL163303 U26595	Homo sapiens Rattus	putative gene containing transmembrane domain	563	78
216	026393	norvegicus	prostaglandin F2a receptor regulatory protein precursor	363	/°
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus	Kupffer cell receptor	567	40
٧٠٠٧) 2003//	musculus	Trabitor our roocher	•••	1
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	636	96
	1.11.021333	norvegicus	kinase		**
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
	1.2.5052.	Trome oupless	11)	***	
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus	*	ļ	
226	AE000218	Escherichia	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
		coli			· _
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus	GTP-binding like protein 2	265	88
		musculus			
229	AF122924	Xenopus	Wnt inhibitory factor-1	316	40
		laevis			1.55
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Human growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific	290	100
	177.50.404	 	phospholipase-D.		100
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
237	X81466	Mus	Embryo Brain Kinase	460	62
220	1164067	musculus	similar to the BPTI/Kunitz family of inhibitors;	284	33
238	U64857	Caenorhabditi	most similar to tissue factor pathway inhibitor	204	33
		s elegans	precursor (TFPI)		
239	AJ250840	Mus	serine/threonine protein kinase	739	63
-57	73230040	musculus	serme an comme protein kindse		"
240	AJ223472	Mus	transcription elongation factor TFIIS.h	222	38
240	7,3223472	musculus	tanscription clongation factor 11 115.11		30
241	Y94906	Homo sapiens	Human secreted protein clone rb649 3 protein	353	52
	15,500	izome supress	sequence SEQ ID NO:18.		[
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-1	591	99
243	L22022	Rattus	orphan transporter v7-3	667	93
		norvegicus		}	
244	AF016191	Rattus	potassium channel	1043	98
		norvegicus	, <u>,</u>	-	1
245	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	645	98
246	Y29868	Homo sapiens	Human secreted protein clone pp325_9.	497	98
247	AF180475	Homo sapiens	Not4-Np	188	83
					99
248	Y17227	Homo sapiens	Human secreted protein (clone yal-1).	690	77

SEQ	Accession	Species	Description	Smith-	%
ID	No.	Į.	}	Waterman	Identity
NO:	<u> </u>	<u> </u>		Score	<u> </u>
		sexta	protein SCLP	 	
250	AF192756	Kaposi's	Orf73	134	34
		sarcoma-	ì	1	ì
		associated		į	
	<u> </u>	herpesvirus			<u> </u>
251	AB022694	Homo sapiens	MOK protein kinase	209	83
252	W55045	Homo sapiens	Neural adhesion molecule (ethb0018f2 product).	469	100
253	L46815	Mus	DNA binding protein Rc	251	67
		musculus		1	!
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus	Citron-K kinase	1201	98
	(musculus		i	Ì.
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841	Oryctolagus	Phospholipase	368	80
		cuniculus	'	j	
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus	L-periaxin	430	72
	1	musculus	•	1	1
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus	SLIT-2	198	40
	1 ,	norvegicus		1	1
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor	636	99
200	1 4 1002	110mo suprems	(GPCR).	1 550	1 33
266	U27269	Mus	sodium glucose cotransporter	204	56
	02.20	musculus	Journal Bross scampoint	1 -0.	150
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus	putative taste receptor TR1	209	39
200	14 12/303	norvegicus		203	"
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus	Fc-gamma receptor	129	26
2,0	170102	pyogenes	1 o-gamman receptor	123	1 20
271	AB009883	Nicotiana	KED	109	26
-/-	1.00000	tabacum		1 .03	20
272	AF137367	Mus	VPS10 domain receptor protein SORCS	899	97
	12 15/30/	musculus	VI DIO GOMMINI IOCOPIOI PIOGONI DOICOD	655	1"
273	L34938	Rattus	ionotropic glutamate receptor	460	86
275	1254750	norvegicus	i ionoropio grammate receptor	. 100] 00
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent	188	74
	1.00000	Tronio sapions	Expressed Protein LIKE PUTATIVE protein)	100	' '
			(isoform 1)	· ·	1
275	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	173	94
213	A 203333	Tionio sapiciis	APOLLON	1 1/3	"
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277	L40380	Homo sapiens	thyroid receptor interactor	430	61
278	AB046851	Homo sapiens	KIAA1631 protein	283	96
279	AC008075	Arabidopsis	Contains PF 00069 Eukaryotic protein kinase	157	43
219	AC008073	thaliana	domain.	157	43
200	1 (02720			1,01	72
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICE1	497	84
283	AF156530	Mus	ETS-domain transcriptional repressor PE1	605	76
204	1 700025	musculus		 	100
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate	647	100
205	3700455		reading frame protein.		ļ
285	Y73402	Homo sapiens	Human secreted protein clone yc25_1 protein	300	90
	<u> </u>	<u> </u>	sequence SEQ ID NO:26.		<u> </u>
286	AF016411	Homo sapiens	KCNA3.1B	137	100
	W89253	Homo sapiens	Human ALP.	688	97
287		Bos taurus	differentiation enhancing factor 1	750	96
288	AF112886				
288 289	AF113131	Homo sapiens	host cell factor homolog LCP	367	44
288			host cell factor homolog LCP plexin-related protein SPA-1 like protein p1294	367 698	100 89

SEQ	Accession	Species	Description	Smith-	%
ID	No.	ļ	1	Waterman	Identity
NO:	<u> </u>	<u> </u>	ļ	Score	
	170000	norvegicus		I	
292	AF102854	Rattus	membrane-associated guanylate kinase-	124	53
202	Vocati	norvegicus	interacting protein 2 Maguin-2	 	
293	X99211	Drosophila melanogaster	ubiquitin-specific protease	143	38
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	185	94
234	124243	Homo Sapiens	sequence SEQ ID NO:92.	183	1 94
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577 1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	182	97
270	174743	110mo sapiciis	sequence SEQ ID NO:92.	102) "
299	B08906	Homo sapiens	Human secreted protein sequence encoded by	605	69
		1 -10-11-0	gene 16 SEQ ID NO:63.	005	} "
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases;	428	72
			Method: conceptual translation supplied by	1	1
	1	1	author	l	
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus	membrane glycoprotein	199	41
	<u> </u>	musculus		1	i
308	AF255614	Rattus	scaffolding protein SLIPR	639	88
		norvegicus			<u>. </u>
309	S79463	Mus sp.	semaphorin homolog=M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium	calcium binding protein	151	36
		discoideum			
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	744	100
212		<u> </u>	124 SEQ ID NO:124.		<u> </u>
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins;	197	38
315	AL021392	- Yrama sanisas	44% similarity to U42767 (PID:g1736918)	020	130
313	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and GENEWISE)	278	38
316	U70209	Mus	polycystic kidney disease 1 protein	165	38
310	0,020	musculus	polycystic kidney disease i protein	103	30
317	AF109643	Rattus	coxsackie-adenovirus-receptor homolog	223	38
		norvegicus	to success and the success of the su	1	1 30
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma	activated protein kinase C receptor homolog	141	38
		vivax	_	1	
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate	232	97
			receptor		
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No.	306	88
	1	1	123	L	
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC	214	97
207	11100000	 	3.1.4.37)	l	<u> </u>
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756	Homo sapiens	protein kinase C mu	540	78
328	G02292	Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
329	AF168990	Homo sapiens	putative GTP-binding protein	877	99
220	S67984	Homo sapiens	anti-HIV gp120 antibody heavy chain variable	581	80
330	li .		region	ſ ·	
	V13016	Uomo sorier-	I DI moneytes related managed (A.A. 104, 4505)	2022	1 00
331	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	2823	98
	X13916 Y87330	Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP-	2823 1127	100
331					

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
	1	 	similarity to P49205 (PID:g1345860)	-	
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	1111	67
336	AF006466	Mus musculus	lymphocyte specific formin related protein	193	75
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia coli	L-idonate transcriptional regulator	928	98
342	D90855	Escherichia coli	glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) chain A, anaerobic	769	99
343	D85613	Escherichia coli	membrane component	399	100
344	M93239	Escherichia coli	transmembrane protein	232	100
345	M60177	Escherichia coli	enterobactin	759	99
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia coli	49 kd protein	1193	96
349	L10328	Escherichia coli	similar to drug resistance translocases	340	90
350	X69942	Mus musculus	enhancer-trap-locus-l	560	82
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+- activated potassium channel	463	80
352	D90777	Escherichia coli	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BibD).	577	100
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVH1	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID NO:258.	165	100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361 362	R28916 U16655	Homo sapiens Rattus norvegicus	Type III procollagen (prior art). phospholipase C delta-4	108 649	65
363	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
364	U47276	Gallus gallus	chicken brain factor-2	104	34
365	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	183	65
366	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
367	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368	AL021366	Homo sapiens	cICK0721Q.3 (Kinesin related protein)	3387	99
369	U70932	Peromyscus leucopus	reverse transcriptase	92	59
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase like	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
375	U49974	Homo sapiens	mariner transposase	172	55

SEQ	Accession	Species	Description	Smith- Waterman	%
ID NO:	No.			Score	Identity
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 6063. Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus	GTP binding protein	1456	91
370	7.52574	musculus	O11 childing protein	1430	1"
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830	Dictyostelium	protein tyrosine kinase	115	44
	1	discoideum			
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228_6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	171	34
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates syndactylus	dopamine receptor D4	105	35
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	Y91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	1047	92
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein; accession number Z21513.	527	78
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus musculus	ADP-ribosylation factor-directed GTPase activating protein isoform b	545	89
405	X92887	Human endogenous retrovirus K	pol/env	162	30
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK022626	Homo sapiens	unnamed protein product	2833	99
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by gene 9 SEQ ID NO:273.	1788	89
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone HTSEV09.	2004	99
411	AB043953	Mus musculus	Chat-H	2628	82
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID NO:148.	1014	92
413	U10542	Pan troglodytes	MHC class I A	265	71
414	AF155097	Homo sapiens	NY-REN-7 antigen	850	95
415	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	88	48
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.	266	89
417	W27651	Homo sapiens	Secreted protein AT205.	481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	3077	87
419	AF255559	Notothenia coriiceps	alpha tubulin	289	68
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
421	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4)))	1446	96
422	AC008075	Arabidopsis	F24J5.4	112	35

SEQ	Accession	Species	Description	Smith-	1%
D D	No.	-,		Waterman	Identity
NO:	1	ł	·	Score	
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1	6268	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 191.	1961	99
426	AB009288	Homo sapiens	N-copine	635	98
427	L12392	Homo sapiens	Huntington's Disease protein	16080	. 99
428	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA- associated protein.	2074	100
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon esculentum	extensin-like protein	613	48
434	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	X73874	Homo sapiens	phosphorylase kinase	3442	97
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type serine protease 1	2630	94
449	U87305	Rattus norvegicus	transmembrane receptor UNC5H1	817	93
450	AF081249	Homo sapiens	JAW1-related protein MRVI1A long isoform	4568	99
451	AC005498	Homo sapiens	R31665_1	316	62
452	M60235	Homo sapiens	granule membrane protein-140	464	73
453	AB036706	Homo sapiens	intelectin	730	88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1 (CIRP-1).	192	67
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by gene 62.	106	40
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
		falciparum			
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	149	43
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	184	54
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein sequence SEQ ID NO:16.	135	47
463	X84960	Triticum aestivum	low molecular weight glutenin	109	33
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	85
465	AF189764	Mus musculus	alpha/beta hydrolase-1	502	59
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by gene 77.	1172	99
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	PI3-kinase	5832	97
470	X70922	Mus musculus	neurotoxin homologue	118	47
471	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75
472	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	72	57

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.	j -		Waterman	Identity
NO:			i e	Score	
			gene 62.	 	
473	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	328	100
474	Y07007	Homo sapiens	Breast cancer associated antigen precursor	1013	97
1		1	sequence.		- '
475	W93254	Homo sapiens	Human ESRP1 protein.	943	80
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	202	60
ĺ			clone HTDAD22	-02	00
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	3427	92
		musculus	g =g =		1
480	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
481	W87701	Homo sapiens	A human membrane fusion protein designated	221	77
-	1	İ .	SYTAX1.	1	' '
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
483	AF210651	Homo sapiens	NAG18	124	59
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
485	G00637	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	70
486	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	149	73
	ļ		3	- "	, ,
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
492	J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
1			3	120	7,
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	197	67
1			clone HTDAD22.	.,,	1 "
495	AC005175	Homo sapiens	R31449 3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis	D4 dopamine receptor	90	48
	1	familiaris	•		"
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935	Peromyscus	reverse transcriptase	213	52
	1	maniculatus	,		"2
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974	116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124_3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135 9.	986	70
507	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	115	33
	,		NO:180.		
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light	184	92
			polypeptide kinase))	.01	7-
509	U43360	Peromyscus	reverse transcriptase	97	62
'		maniculatus		- · · · · · · · · · · · · · · · · · · ·	J2
510	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
511	W79092	Homo sapiens	Human secreted protein dn740_3.	1058	100
512	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	205	64
513	AJ133439	Homo sapiens	GRIP1 protein	2151	100
514	AE003456	Drosophila	CG6393 gene product	259	42 .
		melanogaster	5 COURSE Product	237	74 .
515	Z17206	Xenopus	p46XIEg22	128	40
		laevis	h. Astronom Property	120	+0
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	
517	G03797	Homo sapiens		1766	94
518	AF151083	Homo sapiens	Human secreted protein, SEQ ID NO: 7878. HSPC249	92	40
519	S80864	Homo sapiens		444	98
520	X92485	Plasmodium	cytochrome c-like polypeptide pval	318	50
-20	124703	vivax	PAGE	170	61
		VIVOA		L	

SEQ	Accession	Species	Description	Smith-	1%
D	No.	op. c.		Waterman	Identity
NO:		1		Score	1
521	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	159	59
522	AF121857	Homo sapiens	sorting nexin 7	259	40
523	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	37
524	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	253	73
525	AF119851	Homo sapiens	PRO1722	162	57
526	Y27761	Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
527	G02707	Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
528 529	U47924	Homo sapiens	C8	1112	86
530	G04063 G03203	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 8144. Human secreted protein, SEQ ID NO: 7284.	84	60
531	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	92	65
532	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	75	29
533	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	182	48
534	AF068286	Homo sapiens	HDCMD38P	861	100
535	U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
536	G01955	Homo sapiens	Human secreted protein, SEQ ID NO: 6036.	484	75
537	AF219232	Gallus gallus	qin-induced kinase	206	53
538	AF135022	Homo sapiens	mediator	128	100
539	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	141	59
540	AF016430	Caenorhabditi s elegans	contains similarity to a BR-C/TTK domain	853	39
541	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308)	408	66
542	M29487	Homo sapiens	integrin alpha subunit precursor	517	81
543	AF102530	Mus musculus	olfactory receptor F3	327	73
544	¥73431	Homo sapiens	Human secreted protein clone yb186_1 protein sequence SEQ ID NO:84.	386	100
545	AE004833	Pseudomonas acruginosa	probable TonB-dependent receptor	279	42
546	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	264	53
547	Y69192	Homo sapiens	A human monocyte-macrophage apolipoprotein B receptor protein.	1772	67
548	Y91493	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:166.	176	100
549	G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.	777	99
550	AF044588	Homo sapiens	protein regulating cytokinesis 1; PRC1	1953	88
551	Y29332	Homo sapiens	Human secreted protein clone pe584_2 protein sequence.	1224	94
552	X98330	Homo sapiens	ryanodine receptor 2	24621	99
553 554	Y42782 AB025258	Homo sapiens Mus	Human UC Band #331 protein.	684	95
		musculus	,	501	41
555	AJ010346	Homo sapiens	RING-H2	1468	100
556	W92388	Homo sapiens	Human TR-interacting protein S239a.	538	92
557	AF119851	Homo sapiens	PRO1722	175	59
558	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	183	32
559	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	319	68
560	D86214	Mus musculus	Ca2+ dependent activator protein for secretion	1010	93
561	AF187325	Canis familiaris	melanoma antigen	287	55
562	AJ001981	Homo sapiens	OXAIL	2512	99
563	Z17238	Rattus norvegicus	glutamate receptor subtype delta-1	338	66
564	W30638	Homo sapiens	Partial human 7-transmembrane receptor HAPO167 protein.	371	100
565	AC005620	Homo sapiens	R33590_1	467	97
566	Y99358	Homo sapiens	Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63.	1138	78
567	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	1002	58
568	AF151043	Homo sapiens	HSPC209	798	100

SEQ	Accession	Species	Description	Smith-	%
ID`	No.	•	•	Waterman	Identity
NO:		i		Score	1
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
	Ì	1	sequence.		1
572	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	735	55
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.	254	45
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ	77	70
	1		ID NO:388.		
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis familiaris	D4 dopamine receptor	64	56
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
	V20012	+	Antigen)	225	128
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus musculus	2P1 protein	764	80
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	329	81
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted protein.	110	43
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer polypeptide.	1369	92
593	Y53051	Homo sapiens	Human secreted protein clone dd119_4 protein sequence SEQ ID NO:108.	1112	97
594	Y27658	Homo sapiens	Human secreted protein encoded by gene No. 92.	763	79
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus musculus	COP1 protein	2215	95
597	G03786	Homo sapiens	Human accepted anothin SEO ID NO. 7867	157	65
598'	AF192499		Human secreted protein, SEQ ID NO: 7867.	143	40
298	AF 192499	Mus musculus	putative secreted protein ZSIG37	143	140
599	AF119855	Homo sapiens	PRO1847	236	76
900.	G02872	Homo sapiens		212	73
601	Y00295		Human secreted protein, SEQ ID NO: 6953.	567	88
602		Homo sapiens	Human secreted protein encoded by gene 38. class II cytokine receptor ZCYTOR7	2015	74
	AF184971	Homo sapiens		773	96
603	AF061936	Homo sapiens	diacylglycerol kinase iota dJ963E22.1 (Novel protein similar to NY-REN-2	1333	98
604	AL096828	Homo sapiens	Antigen)	1333	73
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756	Homo sapiens	protein kinase C mu	3915	99
JUU		Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
	1286053		Summa to Difficiation as to be to vitability (0.110.27)	2130	127
607	D86983			1377	90
607 608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	99
607 608 609	W69341 W88627	Homo sapiens Homo sapiens	Secreted protein of clone CG279_1. Secreted protein encoded by gene 94 clone HPMBQ32.	339	82
607 608 609	W69341 W88627 Y27868	Homo sapiens Homo sapiens Homo sapiens	Secreted protein of clone CG279_1. Secreted protein encoded by gene 94 clone HPMBQ32. Human secreted protein encoded by gene No. 107.	339	62
607 608 609 610	W69341 W88627 Y27868 AF202636	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Secreted protein of clone CG279_1. Secreted protein encoded by gene 94 clone HPMBQ32. Human secreted protein encoded by gene No. 107. angiopoietin-like protein PP1158	339 116 2164	82 62 100
607 608 609	W69341 W88627 Y27868	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Secreted protein of clone CG279_1. Secreted protein encoded by gene 94 clone HPMBQ32. Human secreted protein encoded by gene No. 107. angiopoietin-like protein PP1158 PRO0663	339	82 62 100 82
607 608 609 610	W69341 W88627 Y27868 AF202636	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Secreted protein of clone CG279_1. Secreted protein encoded by gene 94 clone HPMBQ32. Human secreted protein encoded by gene No. 107. angiopoietin-like protein PP1158	339 116 2164	82 62 100
607 608 609 610 611 612	W69341 W88627 Y27868 AF202636 AF090944	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Rattus	Secreted protein of clone CG279_1. Secreted protein encoded by gene 94 clone HPMBQ32. Human secreted protein encoded by gene No. 107. angiopoietin-like protein PP1158 PRO0663 Human secreted protein encoded by gene 44	339 116 2164 218	82 62 100 82
607 608 609 610 611 612 613	W69341 W88627 Y27868 AF202636 AF090944 Y02693	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Secreted protein of clone CG279_1. Secreted protein encoded by gene 94 clone HPMBQ32. Human secreted protein encoded by gene No. 107. angiopoietin-like protein PP1158 PRO0663 Human secreted protein encoded by gene 44 clone HTDAD22.	339 116 2164 218 195	82 62 100 82 59

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia	Transmembrane protein dppC	573	90
		coli			
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium bursaria Chlorella virus l	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158 1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	173	78
651	Y32199	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2022379.	1012	100
652	AB032909	Hylobates agilis	dopamine receptor D4	122	32
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens	Human secreted protein encoded by Gene No. 15.	57	37
655	L22455	Rattus norvegicus	mu opioid receptor	116	34
656	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	110	45
657	G02345	Homo sapiens	Human secreted protein, SEQ ID NO: 6426.	459	97
658	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	291	75
659	G02832	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	134	65
660	Y91423	Homo sapiens	Human secreted protein sequence encoded by	333	96
	L	L	gene 11 SEQ ID NO:144.		

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:		1		Score	
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	375	43
			designated HSCOP-6.	i	
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor	480	55
			(rhodopsin family) (olfactory receptor like)		
		J	protein (hs6M1-21))	<u> </u>	
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus	contains transmembrane (TM) region and ATP	757	68
		musculus	binding region		
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus	outer membrane protein	746	81
0,0	12 10,255	norvegicus	outer memorane process	/40	01
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410 5.	261	91
673	G03203	Homo sapiens		106	48
674	AL035587	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	2388	1
675			dJ475N16.4 (KIAA0240)		99
	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus	receptor protein-tyrosine kinase	545	96
		musculus	·	i	1
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus	olfactory receptor	528	77
	İ	musculus		i	
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus	reverse transcriptase	100	37
		maniculatus	,		1 -
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor	2405	99
	1		(MO-REN-46).		
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	183	81
		_	107.		
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	275	75
	1		clone HTDAD22.		1
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- .57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus	torsinA protein	709	45
		musculus			L
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi	similar to Glutaredoxin, Zinc finger, C3HC4	920	51
	<u> </u>	s elegans	type (RING finger)	<u> </u>	l
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

SEQ	Accession	Species	Description	Smith-	T%
ID `	No.	1		Waterman	Identity
NO:	1	<u> </u>	<u>'</u>	Score	1
706	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain	121	95
	1	<u> </u>	ligand (clone 2DD).	<u> </u>	<u> </u>
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
710	M63577	Saccharomyc	SFP1	131	59
711	1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	es cerevisiae		 	<u> </u>
711	AB026291	Rattus	acetoacetyl-CoA synthetase	467	85
712	D21211	norvegicus Homo sapiens	motion to the later (MCD DAS 4 == 2)	379	
713	AF044033	Marmota	protein tyrosine phosphatase (PTP-BAS, type 3)	368 615	44
/13	AT 044033	marmota	olfactory receptor	912	83
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380	100
716	G00577	Homo sapiens	Human secreted protein, SEQ ID NO: 4658.	80	73
717	Y96864	Homo sapiens	SEQ. ID. 37 from WO0034474.	835	99
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid	578	99
			receptor beta4 subunit		
720	AB020598	Homo sapiens	peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	570	74
		1.	designated HSCOP-6.	ļ	!
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma	electrogenic Na+ bicarbonate cotransporter,	111	41
		tigrinum	NBC		L
724	AF127084	Mus	semaphorin cytoplasmic domain-associated	5253	94
	<u> </u>	musculus	protein 3A		
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus	potassium channel	370	100
727	AB029559	norvegicus Rattus	BATI	120	125
121	AB029339	norvegicus	BAII	139	35
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729	AJ011415	Homo sapiens	plexin-B1/SEP receptor	729	56
730	Z93096	Homo sapiens	bK390B3.1 (manic fringe (Drosophila)	142	68
	1		homolog)		**
731	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor	675	99
	1	·	homologue Vanilrep1.		
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia	putative transport protein	592	97
		coli			
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor	2173	99
	1	1	(rhodopsin family) protein similar to high-		}
736	AF132599	Home conier-	affinity lysophosphatidic acid receptor homolog)	245	56
150	AF 132379	Homo sapiens	RANTES factor of late activated T lymphocytes-	245	56
737	X55019	Homo sapiens	acetylcholine receptor delta subunit	883	99
738	X91906	Homo sapiens	voltage-gated chloride ion channel	1978	100
739	AB026116	Homo sapiens	organic anion transporter 4	1444	98
740	D00570	Mus	open reading frame (196 AA)	83	24
	1	musculus	((((((((((((((((((((I	I
	1	musculus			
741	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
741 742	W03626 U66059		Human thyrotropin GPR N-terminal sequence. V_segment translation product	118	40 100
742 743	U66059 AF119815	Homo sapiens	V_segment translation product G-protein-coupled receptor		100 99
742 743 744	U66059 AF119815 X16663	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486)	614 2751 148	100 99 93
742 743	U66059 AF119815	Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486) Human secreted protein encoded by gene 32	614 2751	100 99
742 743 744 745	U66059 AF119815 X16663 W67838	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486) Human secreted protein encoded by gene 32 clone HLTCJ63.	614 2751 148 448	100 99 93 95
742 743 744 745	U66059 AF119815 X16663 W67838	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486) Human secreted protein encoded by gene 32 clone HLTCJ63. Human semaphorin Y.	614 2751 148 448	99 93 95
742 743 744 745	U66059 AF119815 X16663 W67838	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486) Human secreted protein encoded by gene 32 clone HLTCJ63. Human semaphorin Y. Alzheimer's disease protein encoded by DNA	614 2751 148 448	100 99 93 95
742 743 744 745 746 747	U66059 AF119815 X16663 W67838 W57260 W21578	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486) Human secreted protein encoded by gene 32 clone HLTCJ63. Human semaphorin Y. Alzheimer's disease protein encoded by DNA from plasmid pGCS2232.	614 2751 148 448 2414 968	99 93 95 100 65
742 743 744 745	U66059 AF119815 X16663 W67838	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486) Human secreted protein encoded by gene 32 clone HLTCJ63. Human semaphorin Y. Alzheimer's disease protein encoded by DNA from plasmid pGCS2232. Human secreted protein clone yd218_1 protein	614 2751 148 448	99 93 95
742 743 744 745 746 747	U66059 AF119815 X16663 W67838 W57260 W21578	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	V_segment translation product G-protein-coupled receptor haematopoietic lineage cell protein (AA 1-486) Human secreted protein encoded by gene 32 clone HLTCJ63. Human semaphorin Y. Alzheimer's disease protein encoded by DNA from plasmid pGCS2232.	614 2751 148 448 2414 968	99 93 95 100 65

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
751	AB025258	Mus musculus	granuphilin-a	773	41
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	99
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2527	99
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus	vasopressin receptor	979	68
		norvegicus	•		
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8BI92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus musculus	netrin 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus	interleukin-1 delta	309	45
768	AF121975	musculus Mus	odorant receptor S18	268	62
560	47000515	musculus	DDD14	(11	57
769 770	AB008515 Y09945	Homo sapiens Rattus	RanBPM putative integral membrane transport protein	458	50
	1777	norvegicus		700	99
771 772	AF226731 Y27132	Homo sapiens Homo sapiens	AD026 Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	30
784	W49042	Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOV1	1904	91
786	Y91870	Homo sapiens	Human apoptosis related protein.	547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein	1037	100
704	604072	Yama are:	complex component TRAP95	124	62
794 795	G04072 Y69384	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 8153. Amino acid sequence of a 14274 receptor	119	100
704	77/0012		protein.	1260	100
796	W40215	Homo sapiens	Human macrophage antigen.	1358	99

No. No.	SEQ	Accession	Species	Description	Smith-	1%
No.			77-1.12			
AP\$58340 Homo sepiens bepatiocellular curinoma-associated antigen 12 115 99		1	1			
API 98 98 98 98 98 98 98 9		AF258340	Homo sapiens	henatocellular carcinoma-associated antigen 112		99
1999 199863 Homo supjens Human normal uterus fissue derived protein 26. 797 99						
Mono supiess				Human normal uterus tissue derived protein 26	L '''	
B02 D0073 Homo supiens CRI protein. 1913 93 93 97 98 97 97 97 97 97 97						
P92219 Homo sapiens CRI protein.						
Numary N						
Mono Sapiens APP-A receptor preprotein (AA -32 to 1079) 5199 98	002	F72217		CKI protein.	11303) "
W64473	902	V16267	<u> </u>	AND A	6100	-
March Marc						
Bof		-				<u> </u>
808						
808					1	
LAT2						
Homo sapiens	808	AF171669	Homo sapiens		1364	90
Clone HOVBA03 AF108831 Homo sapiens K:Cl cotransporter 3 4561 100	809	W70321	Homo sapiens	Secreted protein CC198_1.	1154	96
AF108831	810	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
Record AF0921372 Homo sapiens Similar to Homo sapiens ribosomal protein L10 Record by GenBank Accession Number L25899	١.	ł	1		1	l
Record AF0921372 Homo sapiens Similar to Homo sapiens ribosomal protein L10 Record by GenBank Accession Number L25899	811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
R12	812	AF092135				
B14 G01563 Homo sapiens Human secreted protein, SEQ ID NO: 5644. 330 100	813	AF283772		similar to Homo sapiens ribosomal protein L10	784	100
L25899					1	
Section	İ	Ì				ļ
Section	814	G01563	Homo sapiens	Human secreted protein, SEO ID NO: 5644.	330	100
816 W95630 Homo sapiens Homo sapiens secreted protein gene clone gn114 1.						
Record R						
Section	*	,	2201110 247.0125		1 300	100
818 AF151800 Homo sapiens CGI-41 protein 1106 95 819 L00352 Homo sapiens low density lipoprotein receptor 3980 100 820 X04434 Homo sapiens IGF1 receptor 5832 99 821 G03844 Homo sapiens Human secreted protein, SEQ ID NO: 7925. 572 100 822 AF212220 Homo sapiens Human glycophosphatidylinositol-anchored 4897 99 823 Y50125 Homo sapiens Human glycophosphatidylinositol-anchored 4897 99 824 AF156778 Homo sapiens ASB-3 protein 2675 98 825 AF096322 Homo sapiens Human secreted protein fragment #2 encoded from gene 28. 1105 100 826 Y07972 Homo sapiens BCL9 5284 96 827 AB032013 Homo sapiens BCL9 5284 96 828 Y13620 Homo sapiens BCL9 5284 96 829 Y91474 Homo sapiens	817	G01082	Homo saniens		549	100
Rightarrow Rig					L	
Record R				L •		
Col						
822 AF212220 Homo sapiens TERA 396 48 823 Y50125 Homo sapiens Human glycophosphatidylinositol-anchored protein GPI-122. 4897 99 824 AF156778 Homo sapiens ASB-3 protein 2675 98 825 AF096322 Homo sapiens neuronal voltage-gated calcium channel gamma-2 subunit 1105 100 826 Y07972 Homo sapiens potassium channel Kv8.1 2435 95 827 AB032013 Homo sapiens potassium channel Kv8.1 2435 95 828 Y13620 Homo sapiens BCL9 5284 96 829 Y91474 Homo sapiens BUman secreted protein sequence encoded by gene 24 SEQ ID NO:147. 87 87 830 X54232 Homo sapiens Human secreted protein sequence encoded by gene 24 SEQ ID NO:147. 87 87 831 X14830 Homo sapiens Human secreted protein sequence encoded by gene 24 SEQ ID NO:147. 87 87 832 Y71262 Homo sapiens Human secreted protein, SEQ ID NO						
R23						
Protein GPI-122. Section 1.25 Pass						
824 AF156778 Homo sapiens ASB-3 protein 2675 98 825 AF096322 Homo sapiens neuronal voltage-gated calcium channel gamma- 2 subunit 1105 100 826 Y07972 Homo sapiens Human secreted protein fragment #2 encoded from gene 28. 1540 100 827 AB032013 Homo sapiens potassium channel Kv8.1 2435 95 828 Y13620 Homo sapiens BCL9 5284 96 829 Y91474 Homo sapiens Human secreted protein sequence encoded by gene 24 SEQ ID NO:147. 98 830 X54232 Homo sapiens glypican 1625 87 831 X14830 Homo sapiens glypican 1625 87 831 X14830 Homo sapiens Human chondromodulin-like protein, Zchm1. 1002 100 833 G03873 Homo sapiens Ry9828_1 1389 93 835 Y38422 Homo sapiens Ry9828_1 1389 93 835 Y38422 Homo sap	٠ دعه	130123	Homo sapiens		4077	39
825 AF096322 Homo sapiens neuronal voltage-gated calcium channel gamma-2 subunit 1105 100 826 Y07972 Homo sapiens Human secreted protein fragment #2 encoded from gene 28. 1540 100 827 AB032013 Homo sapiens potassium channel Kv8.1 2435 95 828 Y13620 Homo sapiens BCL9 5284 96 829 Y91474 Homo sapiens Human secreted protein sequence encoded by gene 24 SEQ ID NO:147. 541 98 830 X54232 Homo sapiens acetylcholine receptor beta-subunit preprotein 2540 100 831 X14830 Homo sapiens Human secreted protein, Zchm1. 1002 100 833 G03873 Homo sapiens Human secreted protein, SEQ ID NO: 7954. 638 96 834 AC030303 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi s elegans dJ1076E17.1 (KIAA0823 protein (continues	824	AE156778	Llomo seniene		2675	00
2 subunit 2 subunit 2 subunit 100 100 100 100						
Record R	023	11030322	Tionio sapiens		1105	100
From gene 28.	826	V07072	Homo conienc		1540	100
827 AB032013 Homo sapiens potassium channel Kv8.1 2435 95 828 Y13620 Homo sapiens BCL.9 5284 96 829 Y91474 Homo sapiens Human secreted protein sequence encoded by gene 24 SEQ ID NO:147. 541 98 830 X54232 Homo sapiens glypican 1625 87 831 X14830 Homo sapiens acetylcholine receptor beta-subunit preprotein 2540 100 832 Y71262 Homo sapiens Human chondromodulin-like protein, Zchm1. 1002 100 833 G03873 Homo sapiens Human secreted protein, SEQ ID NO: 7954. 638 96 834 AC003030 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi s elegans dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 85 36 837 AL121889 Homo sapiens A secreted protein encoded by clone cw1543_3. 1105 67 <	820	10/9/2	Tionio sapiens		1540	100
828 Y13620 Homo sapiens BCL9 5284 96 829 Y91474 Homo sapiens Human secreted protein sequence encoded by gene 24 SEQ ID NO:147. 541 98 830 X54232 Homo sapiens glypican 1625 87 831 X14830 Homo sapiens acetylcholine receptor beta-subunit preprotein 2540 100 832 Y71262 Homo sapiens Human chondromodulin-like protein, Zchm1. 1002 100 833 G03873 Homo sapiens Human secreted protein, SEQ ID NO: 7954. 638 96 834 AC003030 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi selegans dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 98 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein, SEQ ID NO: 4943. 255 92	927	A D022012	Vome comiens		2425	06
R29						
Rene 24 SEQ ID NO:147.						
830 X54232 Homo sapiens glypican 1625 87 831 X14830 Homo sapiens acetylcholine receptor beta-subunit preprotein 2540 100 832 Y71262 Homo sapiens Human chondromodulin-like protein, Zchm1. 1002 100 833 G03873 Homo sapiens Human secreted protein, SEQ ID NO: 7954. 638 96 834 AC003030 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi glycine-rich 85 36 837 AL121889 Homo sapiens dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 98 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543_3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92	627	1914/4	riomo sapiens		541	98
831 X14830 Homo sapiens acetylcholine receptor beta-subunit preprotein 2540 100 832 Y71262 Homo sapiens Human chondromodulin-like protein, Zchm1. 1002 100 833 G03873 Homo sapiens Human secreted protein, SEQ ID NO: 7954. 638 96 834 AC003030 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi s elegans glycine-rich 85 36 837 AL121889 Homo sapiens dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 998 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543_3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein clone yc27_l protein 108	920	V54222	Wama soniana		1626	97
832 Y71262 Homo sapiens Human chondromodulin-like protein, Zchm1. 1002 100 833 G03873 Homo sapiens Human secreted protein, SEQ ID NO: 7954. 638 96 834 AC003030 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi s elegans glycine-rich 85 36 837 AL121889 Homo sapiens dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 998 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543 3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 9		1				
833 G03873 Homo sapiens Human secreted protein, SEQ ID NO: 7954. 638 96 834 AC003030 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi s elegans glycine-rich 85 36 837 AL121889 Homo sapiens dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 998 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543 3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein sequence SEQ ID NO: 114. <						
834 AC003030 Homo sapiens R29828_1 1389 93 835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi s elegans glycine-rich 85 36 837 AL121889 Homo sapiens dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 998 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543_3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td></td<>						
835 Y38422 Homo sapiens Human secreted protein. 964 87 836 U41557 Caenorhabditi s elegans glycine-rich 85 36 837 AL121889 Homo sapiens dJ1076E17.1 (KIAA0823 protein (continues in AL023803)) 998 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543 3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443						
Sample S						
Selegans Selegans						
837 AL 121889 Homo sapiens dJ 1076E17.1 (KIAA0823 protein (continues in AL023803)) 998 75 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543_3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99	830	U4133/		glycine-rich	85	36
AL023803)) 838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543_3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99	027	41 101000		17105/F15 1 (VI) 4 0000	000	
838 AJ011415 Homo sapiens plexin-B1/SEP receptor 1580 60 839 W80398 Homo sapiens A secreted protein encoded by clone cw1543_3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99	837	AL121889	Homo sapiens		998	75
839 W80398 Homo sapiens A secreted protein encoded by clone cw1543 3. 1105 67 840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99	020	A 7011415			1500	L
840 G00862 Homo sapiens Human secreted protein, SEQ ID NO: 4943. 255 92 841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114. 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99						
841 G02650 Homo sapiens Human secreted protein, SEQ ID NO: 6731. 644 97 842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_l protein sequence SEQ ID NO:114. 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99						
842 AF036717 Homo sapiens FGFR signalling adaptor SNT-1 2629 99 843 Y73446 Homo sapiens Human secreted protein clone yc27_l protein sequence SEQ ID NO:114. 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99						
843 Y73446 Homo sapiens Human secreted protein clone yc27_l protein sequence SEQ ID NO:114. 1089 100 844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99						
844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99						
844 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 357 69 845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99	843	¥73446	Homo sapiens		1089	100
845 AF151810 Homo sapiens CGI-52 protein 1443 88 846 X83378 Homo sapiens putative chloride channel 1620 99			L			
846 X83378 Homo sapiens putative chloride channel 1620 99						
847 AC004883 Homo sapiens similar to general transcription factor 21; similar 655 96				putative chloride channel		
	847	AC004883	Homo sapiens	similar to general transcription factor 2I; similar	655	96

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	 		** AF029040 (DVD) -28272072	Score	
848	X99886	Homo sapiens	to AF038969 (PID:g2827207) monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to	963	98
.,	1	110200 020000	P34984 (PID:g464305)] 300	100
850	AB038237	Homo sapiens	G protein-coupled receptor C5L2	1767	100
851	AF124490	Homo sapiens	ARF GTPase-activating protein GIT1	3415	98
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID	1189	99
	<u> </u>	<u> </u>	NO:132.		<u> </u>
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854 855	X17094	Homo sapiens	furin (AA 1-794)	3550	99
800	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone	824	100
	L		HLDRM43.	<u>i</u>	
862 .	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from	388	88
			gene 43.	·	
869	J00123	Homo sapiens	preproenkephalin (1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by	1048	98
871	7.04233		gene 25 SEQ ID NO:305.		
872 ·	L04311 Y29988	Homo sapiens Homo sapiens	GABA-alpha receptor beta-3 subunit Human cytokine family member EF-7 protein.	960	93
873	AF161382	Homo sapiens	HSPC264	1124	99
874	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
875	Y27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	573	96
876	M15530	Homo sapiens	B-cell growth factor	171	56
877	W63681	Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus	neurexophilin	1448	98
		norvegicus			
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted	321	100
880	W88991	Homo sapiens	protein. Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF	528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein	348	100
			sequence SEQ ID NO:106.		•
885	AF070661	Homo sapiens	HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887 888	X92744 Y22496	Homo sapiens	hBD-1	375	100
500	122490	Homo sapiens	Human secreted protein sequence clone cn621 8.	994	94
889	Y41293	Homo sapiens	Human soluble protein ZTMPO-1.	4595 ·	99
890	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	147	63
891	AF208856	Homo sapiens	BM-014	1012	99
892	U29195	Homo sapiens	neuronal pentraxin II	2002	98
893	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein	537	100
DOC .	3461630		sequence SEQ ID NO:34.		
895 896	W61630 M24110	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
897	Z68747	Homo sapiens Homo sapiens	GOS19-2 peptide precursor imogen 38	2018	100 99
898	AF186112	Homo sapiens	neurokinin B-like protein ZNEUROK1	619	100
	AF225420	Homo sapiens	AD025	734	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
904	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688	Homo sapiens	antigen NY-CO-3	771	99
906	AB007836	Homo sapiens	Hic-5	2544	100
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	98
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo I	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162_1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2763	99
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis familiaris	rab8	1064	100
937	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	117	44
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	515	42
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor (PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 463.	667	100
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
950	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

SEQ	Accession	Species	Description	Smith-	1%
ID ID	No.	Species	Description	Waterman	Identity
NO:	140.			Score	lucillity
951	AF110645	Homo saniens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEQ	402	70
132	130111	Tionio sapiciis	ID NO. 496.	102	1,0
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRI	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens	Human secreted protein fg949_3.	1879	100
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202_3 protein	587	100
939		1 -	sequence SEQ ID NO:110.		<u> </u>
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CGI-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene:Acc# AF030433	1466	100
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
968	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
969	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429 `	Homo sapiens	hyperpolarization-activated cyclic nucleotide	6295	100
974	W61619	Homo sapiens	gated cation channel hHCN4 Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657		thyroid hormone receptor-associated protein	3348	99
	J	Homo sapiens	complex component TRAP80		
979	AF044201	Rattus norvegicus	neural membrane protein 35; NMP35	1570	92
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein I	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412-	1553	99
L	<u> </u>	<u> </u>	encoded protein.		<u> </u>
983	Z56281	Homo sapiens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	172	70
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca	juvenile hormone esterase binding protein	226	32
000	C02607	sexta	Live on second and in SEO ID NO. 2229	104	00
989	G03697	Homo sapiens	Human secreted protein, SEQ ID NO: 7778. potassium large conductance calcium-activated	194	100
990	AF204159	Homo sapiens	channel beta 3a subunit	1400	
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi s elegans	VM106R.1	327	40
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	99
994	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
995	AF133845	Homo sapiens	corin	5811	99
996	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	4999	100
997	W62066	Homo sapiens	Human stem cell antigen 2.	284	93
998	Y87173		Human secreted protein sequence SEQ ID	725	100
770		Homo sapiens	NO:212.	123	100
999	Y13379	Homo sapiens	Amino acid sequence of protein PRO263.	1654	99
1000	Y95008	Homo sapiens	Human secreted protein vf3_1, SEQ ID NO:56.	676	47
1001	AF190167	Homo sapiens	membrane associated protein SLP-2	1747	100 .

1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100	SEQ	Accession	Species	Description	Smith-	1%
1003 W73420 Homo sapiens Human secreted protein, SEQ ID NO: 5315. 398 66	ID `	No.			Waterman	Identity
Human secreted protein encoded by Gene No. 2150 100		1		,		
1004 X12791	1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1005 M21323	1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No.	2150	100
M23123		L	1			
		X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1009			Homo sapiens	membrane protein	642	100
DNO.382		X63745	Homo sapiens	KDEL receptor	326	98
Mono supiens	1007	Y35997	Homo sapiens		824	99
Beautiful	1008	AB032918		dopamine receptor D4	92	35
1010 AL150125 Homo sapiens Human secreted protein, SEQ ID NO: 7814. 379 98 1012 Y17531 Homo sapiens Human secreted protein, SEQ ID NO: 4805. 462 100 1014 AF288992 Nagleria gruberi Human secreted protein, SEQ ID NO: 4805. 462 100 1014 AF288992 Nagleria gruberi Human secreted protein, SEQ ID NO: 4805. 462 100 1014 AF288992 Nagleria gruberi Human secreted protein, SEQ ID NO: 4805. 462 100 1014 AF288992 Nagleria gruberi Human secreted protein, SEQ ID NO: 4805. 462 100 1014 AF288992 Nagleria gruberi Human secreted protein, SEQ ID NO: 4805. 462 100 1014 AF288992 Nagleria gruberi Human secreted protein, SEQ ID NO: 4805. 462 100 1014 AF288992 Nagleria gruberi Human secreted protein, SEQ ID NO: 4805. 462 100 1015 AB045292 Homo sapiens M83 protein 1016 AB045292 Homo sapiens M83 protein 1016 AB045292 Homo sapiens M83 protein 1016 AB045292 Homo sapiens M83 protein 1016 H0702632. 1876 100 1010	1009	Y91680	Homo sapiens		1372	99
1012 V17531 Homo sapiens Human secreted protein, SEQ ID NO: 7814. 379 98 98 98 98 98 99 99	1010 ·	AL136125	Homo sapiens		825	98
1012 Y17531 Homo sapiens Human secreted protein Gno BL205 14 protein B18 97	1011					
1014 AF288072 Nessejeria Human secreted protein, SEQ ID NO: 4805. 462 100	1012	Y17531		Human secreted protein clone BL205 14 protein.		
1015 AB045292 Homo sepiens M83 protein 1015 AB045292 Homo sepiens M83 protein 1016 X15940 Homo sepiens M83 protein 1017 Y94873 Homo sepiens Human protein clone HP02632. 1876 100 1017 Y94873 Homo sepiens Human protein clone HP02632. 1876 100 1018 AL024498 Homo sepiens dJ417M14.1 (novel protein) 589 100 1019 X34252 Homo sepiens Lutheran blood group glycoprotein 3054 599 1020 W03516 Homo sepiens Lutheran blood group glycoprotein 3054 599 1020 W03516 Homo sepiens Prostaglandin DP receptor. 1864 100 1021 G03960 Homo sepiens Human secreted protein, SEQ ID NO: 8041. 398 100 1022 Y91689 Homo sepiens Human secreted protein, SEQ ID NO: 8041. 398 100 1022 Y91689 Homo sepiens human secreted protein, SEQ ID NO: 302. 1024 AF132965 Homo sepiens human Secreted protein 1550 100 1025 W92330 Homo sepiens CGI-31 protein 1550 100 1026 R66278 Homo sepiens Therapeutip polypetide from glioblastoma cell 810 1026 R66278 Homo sepiens SIODP calcium-binding protein 476 100 1028 Y41741 Homo sepiens Human PRO704 protein sequence. 1323 100 1029 AJ001014 Homo sepiens Human sercited protein 2. 1354 99 1031 AK023007 Homo sepiens Human sercited protein 766 100 1032 W79900 Homo sepiens Human sercited protein 766 100 1032 W79900 Homo sepiens Human Sercited protein 766 100 1032 W79900 Homo sepiens Human Sercited protein product 766 100 1032 W73473 Homo sepiens Human Sercited protein 752 93 1034 X62433 Homo sepiens Human Sercited protein 1478	1013			Human secreted protein, SEO ID NO: 4805.		
1015 AB043292 Homo sapiens M83 protein 1016 X15940 Homo sapiens ribosomal protein L31 (AA 1-125) 644 100 1017 Y94873 Homo sapiens Human protein clone HP02632. 1876 100 1018 AL024498 Homo sapiens Human protein clone HP02632. 1876 100 1018 AL024498 Homo sapiens Human protein clone HP02632. 1876 100 1019 X34225 Homo sapiens Lutheran blood group glycoprotein 3054 599 1020 W03516 Homo sapiens Lutheran blood group glycoprotein 3054 599 1020 W03516 Homo sapiens Human secreted protein, SEQ ID NO: 8041. 398 100 1021 1022 Y91689 Homo sapiens Human secreted protein, SEQ ID NO: 8041. 398 100 1022 Y91689 Homo sapiens Human secreted protein, SEQ ID NO: 8041. 398 100 1024 AF132965 Homo sapiens CGI-31 protein 1550 100 1024 AF132965 Homo sapiens CGI-31 protein 1550 100 1025 W92380 Homo sapiens CGI-31 protein 1550 100 1026 R66278 Homo sapiens Human TR-interacting protein S103a. 1466 97 1027 X65614 Homo sapiens S100P calcium-binding protein 476 100 1028 Y41741 Homo sapiens RAMP! 806 100 1029 A0001014 Homo sapiens RAMP! 806 100 1030 W63682 Homo sapiens Human REO704 protein sequence. 1323 100 1030 W63682 Homo sapiens Human SR-BI class B scavenger. 2672 99 1033 Y82453 Homo sapiens Human SR-BI class B scavenger. 2672 99 1033 Y82453 Homo sapiens Human SE-BI class B scavenger. 2672 99 1034 Y73473 Homo sapiens Human Secreted protein clone yd178_1 protein 752 93 1036 1037 A1242832 Homo sapiens Muman secreted protein folne yd178_1 protein 752 93 1036 1037 A1242832 Homo sapiens Muman secreted protein polici product 1310 100			Naegleria			
1016 X15940	1015	AB045292		M83 protein	3867	99
1018 AL024498 Homo sapiens Human protein clone HP02632. 1876 100			Homo sapiens			
1018 AL024498 Homo sapiens dJ417M14.1 (novel protein) 589 100				Human protein clone HP02632.		
1019 X83425 Homo sapiens Lutheran blood group glycoprotein 3054 99					,	
1020 W03516 Homo sapiens Homo sapiens Human secreted protein, SEQ ID NO: 8041. 398 100			Homo sapiens	Lutheran blood group glycoprotein		
1021 G03960 Homo sapiens Human secreted protein, SEQ ID NO: 8041. 398 100						
1022 V91689						
December 2015 December 3						
1024		•	•	gene 93 SEQ ID NO:362.		
1025 W92380 Homo sapiens Human TR-interacting protein S103a. 1466 97 1026 R66278 Homo sapiens Therapeutic polypeptide from glioblastoma cell 830 100 1027 X65614 Homo sapiens S100P calcium-binding protein 476 100 1028 Y41741 Homo sapiens Human PRO704 protein sequence. 1323 100 1029 AJ001014 Homo sapiens Human PRO704 protein sequence. 1323 100 1030 W63682 Homo sapiens Human secreted protein 2. 1354 99 1031 AK023007 Homo sapiens Human SR-BI class B scavenger. 2672 99 1032 W97900 Homo sapiens Human SR-BI class B scavenger. 2672 99 1033 Y82453 Homo sapiens Human TGC-440 secretory protein SEQ ID 639 99 1034 Y73473 Homo sapiens Human secreted protein clone yd178 1 protein 752 93 1035 Y86468 Homo sapiens Human gene 48-encoded protein fragment, SEQ 96 90 1036 U09813 Homo sapiens Human gene 48-encoded protein fragment, SEQ 96 90 1037 AJ242832 Homo sapiens calpain 3699 99 1038 X66403 Homo sapiens calpain 3699 99 1039 AJ242730 Homo sapiens calpain 3699 99 1040 AF169968 Mus DNA binding protein DESRT 1453 80 1041 X52563 Bos taurus Dermability increasing protein 383 29 1042 G00368 Homo sapiens Human secreted protein, SEQ ID NO: 4449 75 50 1043 G02532 Homo sapiens Human secreted protein, SEQ ID NO: 6613 60 53 1044 M94582 Homo sapiens Human secreted protein, SEQ ID NO: 6613 60 53 1045 AL080239 Homo sapiens Human secreted protein protein 580 100 1046 AF125101 Homo sapiens Human secreted protein encoded by gene 81 176 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 176 100 1048 AL022238 Homo sapiens Human secreted protein encoded by gene 134 clone 1559 99 1049 W88667 Homo sapiens Human secreted protein encoded by gene 134 clone 1559 99						
1026 R66278 Homo sapiens Therapeutic polypeptide from glioblastoma cell 830 100						
1027 X65614 Homo sapiens S100P calcium-binding protein 476 100 1028				Therapeutic polypeptide from glioblastoma cell		
1028 Y41741 Homo sapiens Human PRO704 protein sequence. 1323 100 1029 AJ001014 Homo sapiens RAMP1 806 100 1030 W63682 Homo sapiens Human secreted protein 2. 1354 99 1031 AK023007 Homo sapiens Human secreted protein product 766 100 1032 W97900 Homo sapiens Human SR-BI class B scavenger. 2672 99 1033 Y82453 Homo sapiens Human TGC-440 secretory protein SEQ ID 639 99 1034 Y73473 Homo sapiens Human secreted protein clone yd178_1 protein 752 93 1035 Y86468 Homo sapiens Human gene 48-encoded protein fragment, SEQ 96 90 1036 U09813 Homo sapiens calpain 3699 99 1037 AJ242832 Homo sapiens calpain 3699 99 1038 X66403 Homo sapiens acetylcholine receptor epsilon subunit CHRNE 2574 100 1039	1027	X65614	Homo saniens	L	476	100
1029 AJ001014 Homo sapiens RAMP1 806 100 1030 W63682 Homo sapiens Human secreted protein 2. 1354 99 1031 AK023007 Homo sapiens unnamed protein product 766 100 1032 W97900 Homo sapiens Human SR-BI class B scavenger. 2672 99 1033 Y82453 Homo sapiens Human TGC-440 secretory protein SEQ ID 639 99 1034 Y73473 Homo sapiens Human secreted protein clone yd178_1 protein 752 93 1035 Y86468 Homo sapiens Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168. 1036 U09813 Homo sapiens Human gene 48-encoded protein fragment, SEQ 96 90 1037 AJ242832 Homo sapiens mitochondrial ATP synthase subunit 9 precursor 698 100 1038 X66403 Homo sapiens acetylcholine receptor epsilon subunit CHRNE 2574 100 1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100 1040 AF169968 Mus						
1030 W63682 Homo sapiens Human secreted protein 2. 1354 99 1031				RAMPI		1
1031						
1032 W97900 Homo sapiens Human SR-BI class B scavenger. 2672 99 1033 Y82453 Homo sapiens Human TGC-440 secretory protein SEQ ID 639 99 1034 Y73473 Homo sapiens Human secreted protein clone yd178_1 protein 752 93 1035 Y86468 Homo sapiens Human gene 48-encoded protein fragment, SEQ 96 90 1036 U09813 Homo sapiens Human gene 48-encoded protein fragment, SEQ 96 90 1037 AJ242832 Homo sapiens acetylcholine receptor epsilon subunit 9 precursor 698 100 1038 X66403 Homo sapiens acetylcholine receptor epsilon subunit CHRNE 2574 100 1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100 1040 AF169968 Mus	1031				1	
1033 Y82453 Homo sapiens Human TGC-440 secretory protein SEQ ID NO:1. 1034 Y73473 Homo sapiens Human secreted protein clone yd178_1 protein 752 93 99 1035 Y86468 Homo sapiens Human gene 48-encoded protein fragment, SEQ 96 90 ID NO:383. 1036 U09813 Homo sapiens mitochondrial ATP synthase subunit 9 precursor 698 100 1037 AJ242832 Homo sapiens calpain 3699 99 1038 X66403 Homo sapiens acetylcholine receptor epsilon subunit CHRNE 2574 100 1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100 1040 AF169968 Mus DNA binding protein DESRT 1453 80 1041 X52563 Bos taurus permability increasing protein 383 29 1042 G00368 Homo sapiens Human secreted protein, SEQ ID NO: 4449. 75 50 1043 G02532 Homo sapiens Human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens Human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens human secreted protein, second labile 1704 50 1046 AF125101 Homo sapiens HSPC040 protein 1560 1704 50 1704	1032					T
1034 Y73473		<u> </u>		Human TGC-440 secretory protein SEQ ID		
1035 Y86468	1034	Y73473	Homo sapiens	Human secreted protein clone yd178 1 protein	752	93
1036 U09813 Homo sapiens mitochondrial ATP synthase subunit 9 precursor 698 100 1037 AJ242832 Homo sapiens calpain 3699 99 1038 X66403 Homo sapiens acetylcholine receptor epsilon subunit CHRNE 2574 100 1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100 1040 AF169968 Mus	1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ	96	90
1037 AJ242832 Homo sapiens calpain 3699 99 1038 X66403 Homo sapiens acetylcholine receptor epsilon subunit CHRNE 2574 100 1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100 1040 AF169968 Mus musculus DNA binding protein DESRT 1453 80 1041 X52563 Bos taurus permability increasing protein 383 29 1042 G00368 Homo sapiens Human secreted protein, SEQ ID NO: 4449. 75 50 1043 G02532 Homo sapiens Human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens interleukin 8 receptor B 1850 100 1045 AL080239 Homo sapiens bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile submit) 50 50 1046 AF125101 Homo sapiens Human secreted protein encoded by gene 81 176 100 1047 W74809 Homo sapiens dJ1042K10.4 (novel protein)	1036	U09813	Homo saniens		698	100
1038 X66403 Homo sapiens acetylcholine receptor epsilon subunit CHRNE 2574 100 1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100 1040 AF169968 Mus						
1039 AJ242730 Homo sapiens polyhomeotic 2 1310 100 1040 AF169968 Mus musculus DNA binding protein DESRT 1453 80 1041 X52563 Bos taurus permability increasing protein 383 29 1042 G00368 Homo sapiens Human secreted protein, SEQ ID NO: 4449. 75 50 1043 G02532 Homo sapiens Human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens interleukin 8 receptor B 1850 100 1045 AL080239 Homo sapiens bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit)) 1704 50 1046 AF125101 Homo sapiens HSPC040 protein 580 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 176 100 1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 cl	1038					
1040 AF169968 Mus musculus DNA binding protein DESRT 1453 80	1039					
1041 X52563 Bos taurus permability increasing protein 383 29 1042 G00368 Homo sapiens Human secreted protein, SEQ ID NO: 4449. 75 50 1043 G02532 Homo sapiens Human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens interleukin 8 receptor B 1850 100 1045 AL080239 Homo sapiens bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit) 1704 50 1046 AP125101 Homo sapiens HSPC040 protein 580 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 176 100 1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone 1559 99	1040		Mus			
1042 G00368 Homo sapiens Human secreted protein, SEQ ID NO: 4449. 75 50 1043 G02532 Homo sapiens Human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens interleukin 8 receptor B 1850 100 1045 AL080239 Homo sapiens bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit)) 1704 50 1046 AF125101 Homo sapiens HSPC040 protein 580 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 clone HMWDN32. 176 100 1048 AL022238 Homo sapiens dT1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone HAIBP89. 1559 99	1041	X52563		permability increasing protein	383	29
1043 G02532 Homo sapiens Human secreted protein, SEQ ID NO: 6613. 60 53 1044 M94582 Homo sapiens interleukin 8 receptor B 1850 100 1045 AL080239 Homo sapiens bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit)) 1704 50 1046 AF125101 Homo sapiens HSPC040 protein 580 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 clone HMWDN32. 176 100 1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone HAIBP89. 1559 99	1042					L
1044 M94582 Homo sapiens interleukin 8 receptor B 1850 100 1045 AL080239 Homo sapiens bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subumit)) 1704 50 1046 AF125101 Homo sapiens HSPC040 protein 580 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 clone HMWDN32. 176 100 1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone HAIBP89. 1559 99	1043			Human secreted protein, SEO ID NO: 6613		
1045 AL080239 Homo sapiens bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit)) 1046 AF125101 Homo sapiens HSPC040 protein 580 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 176 100 1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone 1559 99 HAIBP89.	1044					
1046 AP125101 Homo sapiens HSPC040 protein 580 100 1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 clone HMWDN32. 176 100 1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone HAIBP89. 1559 99	1045			bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile		
1047 W74809 Homo sapiens Human secreted protein encoded by gene 81 clone HMWDN32. 176 100 1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone HAIBP89. 1559 99	1046	AF125101	Homo sapiens		580	100
1048 AL022238 Homo sapiens dJ1042K10.4 (novel protein) 2201 100 1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone HAIBP89. 1559 99	1047			Human secreted protein encoded by gene 81		
1049 W88667 Homo sapiens Secreted protein encoded by gene 134 clone 1559 99 HAIBP89.	1048	AL022238	Homo saniens		2201	100
	1049 ·			Secreted protein encoded by gene 134 clone		
	1050	AF097518	Homo sapiens	liver-specific transporter	2820	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575	100
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.	770	85
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077 1078	L25899 Y91447	Homo sapiens Homo sapiens	ribosomal protein L10 Human secreted protein sequence encoded by	332	76
1078	G01862		gene 48 SEQ ID NO:168.	898	97
1079	AB039723	Homo sapiens	Human secreted protein, SEQ ID NO: 5943. WNT receptor frizzled-3	290 1376	89
1080	AB039723 AB020527	Homo sapiens Homo sapiens	Na/PO4 cotransporter homolog		92
1082	L13802	Homo sapiens	ribosmal protein small subunit	269 499	100 80
1083	W75098	Homo sapiens	Human secreted protein encoded by gene 42	143	81
1083	G03564	Homo sapiens	clone HSXB125. Human secreted protein, SEQ ID NO: 7645.		
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83 88	43
1086	AF090942	Homo sapiens	PRO0657	124	64
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
1091	S72304	Mus sp.	LMW G-protein	146	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123 5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	83	35
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	59
 -	L	1	clone HTDAD22.		L

SEQ	Accession	Species	Description	Smith-	1%
ID.	No.	opas.ac	- Dampaon	Waterman	Identity
NO:		•		Score	1001111
1101	AF119851	Homo sapiens	PRO1722	183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus	ribosomal protein L28	128	69
		musculus	·····	1	1
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor	738	94
1110	AF111108	Mus musculus	transient receptor potential 2	223	79
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by	164	63
L			gene 121.		03
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus laevis	APEG precursor protein	130	40
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 155.	244	97
1121	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	122	65
1122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469	Homo sapiens	Human secreted protein from clone CW795 2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	525	96
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by	542	100
1134	AB017908	Homo sapiens	gene 49.SEQ ID NO:170. 4F2 light chain	2399	02
1134	X51760	Homo sapiens		312	93 55
1136	Y99426	Homo sapiens	zinc finger protein (583 AA) Human PRO1604 (UNO785) amino acid		72
L			sequence SEQ ID NO:308.	917	
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12 (PPRG-12).	623	100
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	AB030235	Canis	D4 dopamine receptor	89	48
L		familiaris			
1144	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7793.	512	85
1148	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
1149	U13642	Caenorhabditi	exon 5 similar to transmembrane domain of S.	247	36

SEQ	Accession	Species	Description	Smith-	7%
ID`	No.	('	[Waterman	Identity
NO:		1		Score	}
		s elegans	cerevisiae zinc resistance protein		
1150	G03438	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1151	G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 5084.	181.	80
1152	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	198	63
1153	X88799	Oryza sativa	DNA binding protein	95	41
1154	D85245	Homo sapiens	TR3beta	155	96
1155	R74272	Homo sapiens	Turnour suppressor protein, p53.	341	87
1156	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	99	41
1			NO:180.	1	1 -
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10	1156	100
	1	i i i i i i i i i i i i i i i i i i i	clone HCEMU42.	1	1.00
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP-	113	31
	\		29 SEQ ID NO:29.	1	
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	149	51
		oup	gene 17.	1	
1168	D90789	Escherichia	Dipeptide transport system permease protein	411	90
		coli	DppC.	ł	
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating	178	59
}		1201110 Suprotis	protein(A49869)	1.70	
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus	tRNA selenocysteine associated protein	249	62
	12.0100	norvegicus	ate at a solutiony stellie associated protein	247	02
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22	107	71
		Tromo suprems	clone HMSJW18.	1	'*
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	69
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7676.	118	46
1187	AB032905	Hylobates	dopamine receptor D4	96	37
''''		concolor	and and the control of the control o	~	"
1188	G00956	Homo sapiens	Human secreted protein, SEO ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7539.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein	187	70
**/*		- TOTHO Subjetts	complex component TRAP230	100	''
1192	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-	202	67
2	1,0-23	TORIO Sapieris	5).	202	37
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 7133. Human secreted protein, SEQ ID NO: 6688.	192	76
1195	W29661	Homo sapiens		2001	
1195			Homo sapiens CI542_2 clone secreted protein.		98
1196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
	X61972	Homo sapiens	macropain subunit iota	149	90
1198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
1199	Y86260	Homo sapiens	Human secreted protein HELHN47, SEQ ID	1089	89
1200	G02602	Warra arriva	NO:175.	ļ.,,,	<u> </u>
1200	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	154	57

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	ļ	<u> </u>		Score	
1201	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
1202	M27826	Homo sapiens	neutral protease large subunit	202	49
1203	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1214	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17 clone HSIEA14.	99	77
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-	725	100
1221	W96745	Homo sapiens	polypeptide. High affinity immunoglobulin E receptor-like	650	98
1222	Y35911	Homo sapiens	protein (IGERB). Extended human secreted protein sequence, SEQ	135	31
1223	Y00278	Homo sapiens	ID NO. 160. Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens		202	95
1225	G01733	Homo sapiens	ribosomal protein S5 Human secreted protein, SEQ ID NO: 5814.	610	100
1227	AF099973	Mus musculus	schlafen2	333	56
1228	G01218	Homo sapiens	Human secreted protein, SEQ ID NO: 5299.	155	81
1229	AF217188	Mus musculus	YIP1B	801	63
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231	X98333	Homo sapiens	organic cation transporter	1704	100
1232	W74955	Homo sapiens	Human secreted protein encoded by gene 77	212	53
1233	Y94940	Homo sapiens	Clone HOEAS24. Human secreted protein clone yi62_1 protein	526	100
1234	U76618	Mus	sequence SEQ ID NO:86. N-RAP	482	82
1025	AE044004	musculus	L-12	200	07
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens	Homo sapiens CH27_1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane-associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
1245	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
1247	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ	168	39

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	ŀ		Waterman	Identity
NO:	<u></u>		<u> </u>	Score	
			ID NO. 160.		
1248	AF072509	Rattus	glutamate receptor interacting protein 2	559	90
		norvegicus	<u> </u>		1
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by	1087	97
	<u> </u>		gene 27 SEQ ID NO:131.		ľ
1251	L15313	Caenorhabditi	putative	858	59
		s elegans		1	1
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate	278	75
	<u> </u>		reading frame protein.		1
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-1	222	54
1257	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded	81	94
	1	1	from gene 26.	l	
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	986	100
	1		ligand (clone 3TW).	1	1
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor	288	71
			sequence.	1	1
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ	723	93
			ID NO:2.		
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus	phosphatidylinositol 5-phosphate 4-kinase	859	95
	1	norvegicus	gamma	1	
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus	LMBR2	552	76
	1 .	musculus		1	l '-
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras	820	98
	J	į	GTPase-activating protein p135 SynGAP)]	
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform .	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme	1280	100
	j	j	similar to acetyl-coenzyme A synthethase	•	
_		}	(acetate-coA ligase))		l
1274	AF064748	Mus	S3-12	3523	61
		musculus			
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	Y30715	Homo sapiens	Amino acid sequence of a human secreted	643	90
		1	protein.		1
1277	AF146760	Homo sapiens	septin 2-like cell division control protein	707	100
1278	Y05069	Homo sapiens	Human PIGR-2 protein sequence.	281	46
1279	X59668	Oryctolagus	aorta CNG channel (rACNG)	267	85
	L	cuniculus	, ,	-	
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1	1635	100
1283	AF117814	Mus	odd-skipped related 1 protein	357	98
	1	musculus	· · · · · · · · · · · · · · · · · · ·		[
1284	U87318	Xenopus	NaDC-2	535	60
	1	laevis			
1285	AF061346	Mus	Edp1 protein	452	68
	1	musculus	• •		
1286	AB030182	Mus	contains transmembrane (TM) region	582	68
		musculus	(11.4) 1081011	JUL .	ا
1287	A13595	synthetic	immunosuppresive protein PP15	185	97
		construct		100	71
	1 200 00 00 00 00	Homo sapiens	ser/arg-rich pre-mRNA splicing factor SR-A1	837	100
1288	AF254411				100
1288 1289	AF254411 AF084205	Rattus	serine/threonine protein kinase TAO1	319	98

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	Jopania	2000 Q 200	Waterman	Identity
NO:		1		Score	1 dominy
1290	AF038563	Homo sapiens	membrane associated guanylate kinase 2	523	100
1291	AF034837	Homo sapiens	double-stranded RNA specific adenosine	468	100
	14 05-057	Tiblio supiciis	deaminase	400	100
1292	M15888	Bos taurus	endozepine-related protein precursor	937	87
1293	AB010692	Arabidopsis			
1273	AB010092	thaliana	ATP-dependent RNA helicase-like protein	636	45
1204	177000000			1	L
1294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
1295	W67828	Homo sapiens	Human secreted protein encoded by gene 22	504	98
	<u> </u>	<u> </u>	clone HFEAF41.		
1296	AC004832	Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
1298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
1300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1302	M77693	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	100
1302	G01331	Homo sapiens			96
1303	G01491		Human secreted protein, SEQ ID NO: 5412.	254	69
	AF148509	Homo sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
1305		Homo sapiens	alpha 1,2-mannosidase	602	98
1306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying protein SEQ ID NO:1.	332	98
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
1309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus musculus	arsenite inducible RNA associated protein	688	70
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44	807	97
1316	A 1070078	TT	clone HE8CJ26.	 	100
	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1317	AB041533	Homo sapiens	sperm antigen	2607	98
1318	U19617	Mus musculus	Elf-1	806	92
1319	U82598	Escherichia coli	ferric enterobactin transport protein	768	100
1320	D90892	Escherichia coli	SORBITOL-6-PHOSPHATE 2- DEHYDROGENASE (EC 1.1.1.140) (GLUCITOL-6- PHOSPHATE DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE).	709	100
1321	W67847	Homo sapiens	Human secreted protein encoded by gene 41 clone HPBCJ74.	601	92
322	AJ276101	Homo sapiens	GPRC5B protein	466	93
323	AJ276101	Homo sapiens	GPRC5B protein	504	97
324	Y58628	Homo sapiens	Protein regulating gene expression PRGE-21.	1584	100
325	U91561	Rattus	pyridoxine 5'-phosphate oxidase	1277	89
-	[norvegicus	Ly	1	ر ا
326	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
327	Y32206	Homo sapiens	Human receptor molecule (REC) encoded by	1531	90
220	AE161040	 	Incyte clone 2825826.	655	05
328	AF151048	Homo sapiens	HSPC214	657	85
329	Y10530	Homo sapiens	olfactory receptor	1645	100
330	AF180681	Homo sapiens	guanine nucleotide exchange factor	4314	99
331	AF111856	Homo sapiens	sodium dependent phosphate transporter isoform NaPi-3b	3591	99
	(.				
332	Y13583	Homo sapiens	G-protein coupled receptor	2171	100

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339 ·	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

TABLE 3

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	I Amino out a comment (AmAlonia Color :
NO: of	NO: of	hod	ID NO:			Amino acid sequence (A=Alanine C=Cysteine,
nucl-		nou		beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
]			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]]	i	peptide		/=possible nucleotide deletion, \=possible
Ĺ				sequence	Í	nucleotide insertion
1	1351	A	2	337	1	TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC
1	i	1	ł	•	{	HWPQAPHRA***GLLPPRWLGHGLPGGPAAP
	l		Į			WAASOWVDGVAGRLPGPAWSWHASGAAPA
ŀ	l	ł	ŀ	}	}	QPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL
ļ.						QIETT\YHHTPIRMAKIOKT/GHHOC**ECGAT
			i			GTLIHGWWGCKVVEPLGKTVWQIPK
3	1353	A	40	3	314	· HASAHASVVLKDNSELEQQLGATGAYRARA
] ~	1555	**	1 **	"	317	LELEAEVAEMROMLOLEHPFVNGADKLRPD
1	ľ	ľ	1		ł	SMYVHLNEL*OSLVENMLLTVVDTH\RTPI*R
i	ļ					SCNYTLALILFL
4	1354	A	74	2	292	
7	1334	Α	′*	2	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR
l	Ì	ł				QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT
	. '					VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH
<u> </u>						VLPLP
5	1355	A	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP
j ,		j				NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA
						FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI
			l I			AGMLGAVISGIWLDRSKTYKETTLVVYIMDT
i						GGAWWCYTFYLGTGDTCG*CFITAG\TMGFF
		,	ł l			MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA
			i			OVFGIIFTISOGOIIDNYGTKPGNIFLCVFLTLG
l .]			AALTAFIKADLRROKANKETLEN
6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A
,			-		-	YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE
						*CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNOLLVWYP
'	2337		[]	- 1	0,2	ATALADNKPVAPDRRISGHVGUFSMSYLESK
			Ì			GLLATASEDRSVRIWKGGDLRVPGGRVQNIG
i ·		1				
L		لـــــــــــــــــــــــــــــــــــــ	L			HCFGHSARVWQVKLLENYLISAGEDCVCLV

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
Seq-	uence	[09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l dance	<u>}</u>	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide	Sequence	/=possible nucleotide deletion, \=possible
1		l	Ï		1	nucleotide insertion
<u> </u>			├──	sequence		
		İ		İ		WSHEGEILQAFRGHQGRGIRAIAAHERQAWV
Į.		l	İ	l	l	ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ
İ		j		1		VP**ARYTQGCDSGWLLATAGSD*YRGPVSL
1]]]		j	*RRGQVLGAAARG*TFPVLLPAGGSSWSRGL
	1	İ				RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS
<u> </u>		L		<u></u>		WEGAQLELGPAWL
8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF
	ļ					LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL
1))]	j	LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV
			L			QCLGFVDSDSRKMVSTLT
9	1359	Α	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN
_	<u> </u>			_		KSSEFNEGPERERMDV
10	1360	A	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY
		1	1			FEEVQRLRFEVHDISSNHNGLKEADFLGGME
1		ł	l		ŀ	CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA
1	l		1		1	EELSGNDDYVELAFNARKLDDKDFFSKSDPF
	1					LEIFRMNDDATQQLVHRTEVVMNNLSPAWK
1					[SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK
	ĺ				}	HDFIGEFTSTFKEMRGAMEGKQVQWECINPK
		1	1		}	YKAKKKNYKNSGTVILNLCKIHKMHSFLDYI
1					ł	MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP
1	ĺ	ĺ	ľ			YQPNEYLKALVAVGEICQDYDSDKMFPAFGF
]	ļ	l			ļ	GARIPPEYTDSHDFAINFNEDNPECAGIOGVV
		ĺ	'			EAYQSCF\PKAPTFTGPTNICPHSSRKVAKFRR
1		1				SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP
l i		ł	i .			DNPGGHFV
11	1361	A	147	614	9	ACARKOLLGRTVFIWFVGOLLGGELKGYSKT
••	1501	'`		014) 1	NTTSSRPASSRG\TLSSSSSSSSLTKDALPSSL
						KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI
	ì	l		į		KILLSSSVKYLE*KRTSCCFPDSSESKLSOLSS
			ļ			DERVSMGTSSRKPTNSSSSLGALKMSATS*G
·						SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS
						TILTRDC
12	1362	Α	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI
12	1302	A	1//	12	410	
						DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM
						VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA
	· .				· ·	ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI
12	12/2		240		100	NLLTMNVY
13	1363	Α	249	535	105	WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ
						MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS
					,	TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI
						DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC
	_					QVLGTPKKVSTLVPKLL
14	1364	Α	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH
						FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT
						*SFFFIFSWGTNGCLLSAITYACYAAICHPLLS
						TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM
ļ						AIEFLLECDONIT/KLICENT*KNIAKNI*KRRV
				l		TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI
						KOTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK
		••				SKDVFSKPVNIFWALEESVLGVKARQPKPFFA
				J		AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV
, ,				1		GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	A	298	68	208	RKRTNPIKLDKKFEHFKNEDI*ITSKHTKMW
. *′	130/	Δ.	270	00	200	
18	1369		300	904		VSSLAMKEMLTKTTM
10	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL
		لـــــــــــــــــــــــــــــــــــــ	1			ELLTSGDPPALASQSAGITGMSHCARPKGHFG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR WK*DY*C/LQEVTDPIMEKGKKKKTASFFK GQPHQSTNALLRRCVR*RYHLS/TVETAGLP*KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR ISLMSSWDYRRPPQ
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS WICRLRPLLWRAVREYLSKLKNAFLSFDPGV SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA AV*NKPRHILLSHIWKDVQNILLK
20	1370	A	304		1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA PLHSSLGNTVKP*PKNQKPKQNRSRHGQ\GF MAGQGQSRPAAR*PPCPALTPASHSAGTWPP RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR CRALPGRLCSAPAAGLRRARPRLSESRRGNSP PASPAAASARCPSWGPSCPARPPSRPAAGTEP AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW
21	1371	A	326	799	1587	ALVRSRGG GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP LRHVRLFSAGAPRGAATPCPPALLHGPAWPP ARPMFRGHPPVRPLGPWGKVAAGPRALCLA GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA AAQAEPGADPEPEDKDQAAESRPAGAMSLSA QGSGPVGGOGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP GAPCYPGHPHLENPHLEHLLTWRTVTWSTLL PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP GTVVSP
23	1373	A	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL NNEKRKMKKRKEEKKKCRERMQRRSKWRR EEKKE*RREE\EERKKEKEDRKERRKETSPRG SRRLLRD
24	1374	A	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP WEGRTTAQWSLHRKRHLARTLLVSRVRGPO
25	1375	A	384	373	128	YLITTILETGYLWKNRHSDQ*KRTENPERDQH KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD KKINLNLKPHTKLTPNIKKN
26	1376	A	397	383	165	EVKNTNPFIFSGTNLTTWIRSI*RKSDEINQRTK *MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK IVP\FNGSKNKIAFINLTKYQNIQNRHAENYKI LVNKIEDLNKWRNVLLSWIGRRNIINTMT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ļ		[amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ì		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL
i i				1		ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF
ļ		1				YQTFKEEL/II/ILHKLFQTIKYGRILPNSVYETSI
1		,				TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA
20	1270		43.4	205	100	NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK
						VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID
-20	1200	_	166			PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS
i i						SWDYRYAPPRP\ANF*FLVETGFYYVAQAGL
31	1381	A	462	393	2	KLLSPGDLPALAS
31	1361	A	402	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI
						CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV
					i .	EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK*
						KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK IFAN
32	1382	Α	474	125	471	VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK
32	1302	^	7/7	دعد	7/1	EIKAYT/RRICTPMFAAPVSVIA/RN*KOSK/CO
1						KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT
ł						ILRETDRIHKTTYDVISLI
33	1383	A	488	1825	2	KSACSFICSEEOPASPSPLKPGTYASET\RPRDP
"	1363	7.	700	1025	2	HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT
						PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR
						APGGSAGS*\GLPSAGGSRGRKGWRAAGROP
]						STR*GRPGRHGGRGE*AGHPEPROSALOSAG
				į		L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P
						PSVLSRS\GS*G*AASGTASSPRSHSSRLGPP
1						SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG
1						AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA
						SPQTAAGAGSPVQWALSRATG*TGETGSWC
1						AGGTHQATHLTAAWVCPPTWSVRPGGSGPA
						AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP
			J			SPASSEVALSSGSCWPDQAPGPARGSPPAPLA
						PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL
						SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP
				İ		L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS*
					!	GGRSPAGTGHLGAOTVASPH*GHWPTALSCL
						WASASPPGPEAPPQTGACIGTNCRYRAASAR
				'		RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER
j						GALTHRPRAPDE
34	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E
				İ	İ	AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA
	ļ]	l		RLS\PPLASCGGRGPPGGAACATCAPPAGPAR
	į	' I	· 1		1	SSRCRRRSPPE*GPR*PSRPARPSPGSAASRRQ
ļ		ļ		-		KLTPCRCQFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN
	Į	ļ	1		İ	LTELVVAVTDENIVGLFAALLAERRVLLTAS
		ļ	J	ļ		KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH
		_	_	.]	Ì	LLDYC*CPPLPRT
36	1386	Α	512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA
]		- ([FLGLAAGGQTLCPAGELPGHARAQASGAPGS
] [ſ	Ì	İ	[ĺ	VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL
]]			l	}	ļ	GHTRPARPRPV\PFAPAVPQEPGGQGHGAA/P
1	ł	1	ł	l	ł	PATGHSAPRGCPPARAAPTGSATPAPPPAACA
						AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT
	}	- 1		ļ	į	PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA
[[ļ		GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP
<u> </u>	1					VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR
[Í	i	j	ľ	AAVARRLRSWNACGLSRVAGRSSASYPGRE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	F=rnenylalanine, G=Glycine, F=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	"""		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
100		1	7.7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}		}	[peptide		/=possible nucleotide deletion. \=possible
			j	sequence	}	nucleotide insertion
						GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD
1		ŀ)		j	GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS
'					·	VSAAPQSPRTRCPRGCAAAAGLCVLAAAGAS
1						HGA\GLPGVRVHTQRVHIH*GAG/GCQTPRPR
1			1 1			LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA
			ļ i	,		ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP
i .			1			NHQAVGLEASGALQAGHRDELPTMVQLLDH
L						SPDYPLKGRPHAP
37	1387	Α	620	828	1	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
						WEASPEMQSKCHQKGKNNQTECFNHVRFLQ
				'		RLNSTHLYACGTHAFQPLCAAIDAEAFTLPTS
						FEEGKEKCPYDPARGFTGLIIDGGLYTATRYE
1						FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE
		Ì				AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH
						SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH
1					1	\DAEFVFSVLVRESKASAVGDDDKVYYFFTE
20	1200				400	RATEKESGSFTQSRSSHRVARGIPPL
38	1388	Α	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG
						TSEEGKEG*EVPV*LPVSPPLPRPLQKMLDYL
]						KDKKEVGFFQSIQALMQTC\GEKVMADDEFT
1						QDLFRFLQLLCEGHNNDFQNYLRTQTGNTTT
39	1389	A	767	1	1030	INIIICTVDYLLRLQESI
37	1307	^	/6/	1	1030	TLDLTGPLLLGGVPNVPKDFRGRNRQFGGCM
						RNLSVDGKNVDMAGFIANNGTREGCAARRN FCDGRRRQNGGTCVNRWNMYLCECPLRFGG
	'					KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
						VRGLHIQVRQPLVVYAAFTVDSHRPLOETVL
1 1			l			RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
						ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
1 1						PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG
1						LRVTDGEWHHLLIELKNVKEDSEMKHLVTM
1						TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA
						SEDKVSVRRGFRGCMQVRGGCGGRGEACPS
1						QAAPRL
40	1390	A	801	69	399	IHKIIIHKEDLNKWKYILCSGMERLSTVMIPVV
						PQIIYKFNA*Q\VILKFTW*E*GAKITILRKNKL
						RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR
		<u>.</u>		<u> </u>		VNKSVVLVQVTIM
41	1391	Α .	835	7	195	SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD
]				SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT
42	1392	A	841	1	415	GSTHASGYDKTPDFILQVPVAVEGHIIHWIES
]		1		1	1	KASFGDECSHHAYLHDQFWSYWNSLKHRTW
] [ļ]	ļ	ļ	QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG
1		1		1	ļ	FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL
<u> </u>						KACFPTNIVTL
43	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG
1		i	1		Į	LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP
<u> </u>						QRPMLPPSHAGLARPPPPEPISVP
44	1394	A	853	452	1	LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK
'		j	j	j		SPRCIPRTSLYFTICC/PPALQL/SPIEDPPATYRS
]			ļ	}		PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG
[l	ļ			QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR
<u> </u>	100-					PLTFSTRRNVDPEIPERFR
45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG
; l	. 1	ł	İ	,		QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT
<u> </u>	1206	 -l				WLSMSMGK
46	1396	A	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL
	1	ł	ł	Į.		EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR
L		1				VFLFHQLNIT**CLHFFTMTTFIAIPFSFLFLGR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914 944 963 967	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion D/KSLAMLPRLVSNSWPQVILPP QLQNLASRGCL*SQLLRRLRENRLNPGGGG CSEIAP\CTPAWVTQRDFFRKKK HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC PRKRESWWGERLP/PRGFPPAAEDAPAPGWK GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP
50	1400	A	973	45	421	RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAQP**TPRCPAALRAGAHIGRVGRPY EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE
51	1401	A	992	2095	194	E IRRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKGLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPTTAASSMTAASSTPMILALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAQGPISQVSVDQPVV NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV *LMLA*TKDPSNRPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT OKRAA\LYTWHVLEQLEILRQINQQSHGPG
56	1406	A	1044	5	429	SVLTLQTRSPSKPLSRKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of- peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV TCLRRREKDGVIVDVLSDTASNHNGFPVEEH ADDTHPARLQGPTLRSQPMGPLKHKAFEERA NLGLVQRRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL PCLGCPTXATCRLYQTTVAVVF
59	1409	A	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL LLLAVQQSCLADHLLTASWGGK/DPIPTKALG EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCOPFHLVMRDLLQLGQDIPQGCHY LEENHLIHRDIAARNCLLSCAAPTRAATIGDF GMARYIYRTRYYQLGDRAL/LPRK WMPPEAL LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR TN
	1412	A	1080	1	859	VVEFLWSRRPSGSSDPRPRRPASKCQMMEER ANLMHMMKLSIKVLLQSALSLGRSLDADHA PLQQFFVVMEHCLKHGLKVKKSFIGQNKSFF GPLELVEKLCPEASDIATSVRNLPELKTAVGR GRAWLYLALMQKKLADYLKVLIDNKHLLSE FYEPEALMMEEEGMVIVGLLVGLNVLDANL\ CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE HERITDVLDQKNYVEELNRHLSCTVGDLQTK IDGLEKTNSKLQERVSAATDRICSLQEEQQQL REQNELIR
63	1413	A	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI HTGEKPYTCGECGKTFRQSANLYAHKKIHTG EKPYTCGDCGKTFRQSANLYAHKKIHTGEKP YKCKECGKAFKSYYSILKHKRTHTRGMSYEG DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE KAFNHTSICCRHKKN
64	1414	A	1084	946	1	KKQDLSSSLTDDSKNAQAPLALTESHLATLA SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS FFSWLTTGLTTQQRTAIE\NATVAFFLQCI\SC HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI S\GFIRRLFLQLMLEDEKVTMFLQSPCPLYKG RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415.	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA SVALHKLSNALV
66	1416	A	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD TLPVAAAFTETVNAYFKGADPSKCIVKITGE MVLSFPAGITRHFANNPSPAALTFRVINFSRLE HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL MTHLK
67	1417	Α	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA PYYFLLDLCCSDILRSAICFPFVFNSVKNGST WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	noa	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	иелсе	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			· ·	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	!			peptide	1	/=possible nucleotide deletion, \=possible
Ĺ	Ĺ	l		sequence		nucleotide insertion
						RYL
68	1418	A	1106	1 .	1326	MGKISATGINMGTKCSWALVWHLESYDPKH
				1		YEREGMQDWKTASGQSEEATQQSSQKPQPH
		ļ				YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF
i		ļ				PGRSRARPPRTRQQRRGAAAGPGRGAVRLG
						HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA RNSDASGPASLSRTLGRASSPRPPQAPDVTAP
1	·	[i 1		ł	SPAALAPRAARGGSRAAALAGAEAEEPLRTL
Ţ			(APRPTRAAAPPPPPPPPPPLPPGAPPPPVRCVSR
						RARAPPWR/PAATGPPP\RPVAPSRKLGSARAP
			1			APALQIRKGTSSGLPGRGGGSGPGNNLSSVA
			1			GNWRGSSFAVERPGMAKYQGEVQSLKLDDD
]			SVIEGVSDQVLVAVVVSFALIATLVYALFRNV
			,			HQNIHPENQELVRVLREQLQTEQDAPAATRQ
		i]			QFYTDMYCPICLHQASFPVETNCGHLFCGSLT
						PNSIW
69	1419	A	1107	2	466	FDTARLHEFGTSITQIFAVDNREDLQKWMEA
			1			FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF
1						LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE
		l l		'		TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS
70	1420	A	1111	698	23	RKRY/SYPASEPLHDEKGKKRQAPLPPSDK ALRRLHYVRATKV\FLSFRPFWREEHIEGGH
/ .	1420	A	1111	096	23	SNTDRPSRMIFYPPPREGALLLASYTWSDAAA
						AFAGLSREEALRLALDDVAALHGPVVROLW
				'		DGTGVVKRWAEDQHSQGGFVVQPPALWQT
						EKDDWTVPYGRIYFAGEHTAYPHGWVETAV
İ						KSALRAAIKINSRKGPASDTASPEGHASDMEG
						QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL
						QNTTHTRTSH
71	1421	A	1119	2	385	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE
,						PPGPPEQAGLSQFHLEPETQNPETTEEIQSS\LQ
						QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL
						EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	A	1127	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT
					•	GQKVAIKIVNREKLSESVLMKVEREIAIL\RLI
	<u>'</u>					EHPHVLKLHGVYENKKYFPPDELTSGPSMLA
						QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG GELFDYLVKKGRLTPKEARKFFRQIVSALDFC
						HSYSICHROLKPENLLLDEKNNIRIADFGMAS
						LQVGDSLLETSCGSPHYACPEVIKGEKYDGR
				i		RADMWSCGVILFALLVGALPFDDDNLRQLLE
						KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR
						LSLEQIQKHPWYLGGNFIS
73	1423	Α	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV
}	·					FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE
	1		1	' I		MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG
					1	SDFLCTEWKASNSVPTSVHQLRPADIKVVAA
		ļ		l		LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG
}	}	ł		Į		GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG
		ŀ				TAGLNVAAEGARARDMPAQAWDLVERMKN SPDINLEKDWKLVTLFIGGNDLCHYCENPEA
				ļ		HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHOPLREASWLA/LPPIGLWGT
'	4747	^	1137	-	TOV	DSPLCCVEVAIPCNKGAHSVGLKGWLLAOG
'	1	į		[]	VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE
	1	- 1		ļ	}	EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE
					ļ	AADPAPVHTTAHPKGA
75	1425	A	1147	2	413	PFPHQHPQEP\KGSCWPQSALRGQCPGPVLGV
<u> -</u>	}	ļ		j	l	TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR

SEQ ID NO: of	SEQ ID	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	NO: of peptide	100	in NO:	beginning nucleotide	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	цепсе		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	}	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	}	1	1	peptide	ŀ	/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
						RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ
	1			ŀ	}	DDESGQKKLHGLQAILVHEASGTTAITATAT
126	1406	<u> </u>	11.55			GYQESHLSSAR
76	1426	A	1155	38	410	PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK
	ļ	l				PDCKEIWIFWWGDEPNLVVQYIMNCMLWK
]				KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF
j		1				T
77	1427	A	1162	526	350	RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV
						LRGNLLRVFRQVEKVQEENKWQSPLED
78	1428	A	1171	1	1293	MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP
		(1		SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT
						TSPRNWCIKMVCNPWFECVSMLVILLNCVTL
1		ł				GMYQPCDDMDCLSDRCKILQVFDDFIFIFFA
1 1		1				MEMVLKMVALGIFGKKCYLGDTWNRLDFFI
}						VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA
						INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF
]						FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL PP\YYQPEEDDEMPFICSLSGDNGIMGCHEIPP
)						LKEOGRECCLSKDDVYDFGAERODLNASGL
						CVNWNRYYNVCRTGSANPHKGAINFDNIGY
	·	}				AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI
				İ		YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP
						GVAAESLLLRGWVLWLPGGG
79	1429	Α	1175	1	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN
1		1			İ	FLSATHLGGLFPPWPLVEERKLKPKASQQCPI
.						CHKVIMGAGKLPRHMRTHTGEKPYMCTICE
1 1						VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF
80	1430	Α	1182	25	198	VHNYDLKNHMR EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT
"	1450	^	1102	2.7	170	PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	Α	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS
}						AIWQQAREVVRFNGLEDRVHVLPGPVETVEL
1. 1			l			PEQVDAIVSEWMGYGLLHESMLSSVLHARTK
						VVKDGGFFLPXSSELFM
82	1432	A	1187	2 .	716	DFVDAARNLPLESTKSPAEPSKSVPSLENDPRA
'			·			SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT
			-		. !	SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH GREATER OF A HELOPET ATTEMPT OF THE PROPERTY OF THE PROPERT
			Ì			GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD
			,	}		TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP
]					j	SQESNVSLSGSSRSALFERDDHGKAEAPSPSF
		-	-			DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC
[WGRGHGCGQEALSTSHGYHLFCALLTGFLFA
					Į	SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF
لـــــا			l			Q
84	1434	A	1192	45	476	LGDVGFWVERTPVHEAAQRGESLQLQQLIES
		l	. 1	[Ì	GACVNQVTVDSITPLHAASLQGQARCVQLLL
	ļ	ł	l	l		AAGAQVDARNIDGSTPLCECLRLGQHRVCEA
		1	ł			LAVLRGQGQPSPVHSVPPARGLHXREFRMC*
85	1435	A	1194	69	410	GFLFDVGXNLEAHEFHFGEP
ره	1433	Λ	1174	לט	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR SCLEARKSOPDEKLLSALHNSRTWN*EPRRSO
	Į	- 1)	ļ	•	HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT
.	Ì	Į			1	GRSPCPSLPGTTRTNSLL
86	1436	A	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC
'		·]		·		NLGFFLEGHAVLTCHAGSENSATWDFPLPSC
		j				RADDACGGTLRG/AEWHHLQPPLPLG/ATKN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	J	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	[USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	(ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}			l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
•		1	1	peptide]	/-possible nucleotide deletion, \-possible
		<u> </u>	 	sequence	ļ	nucleotide insertion
		}	1		!	NADCTWTILAELGDTIALVFIDFQLEDGYDFL
107	143-		1215		L	EVTGTEGSSLW
87	1437	A	1216	226	964	GTARFGPMVGFGANRRAGRLPSLVLGVLLV
		ĺ	Ì			VIVVLAFNYWSISSRHVLLQEEVAELQGQVQ
1			ĺ		i	RTEVARGRLEKRNSDLFAVVGHAQETDRPEG
					1	GRLRPPQQPAAGQRGPREEMEDDKVKLQNN
			l		ł	ISYQMADIHHLKEQLAELRQEFLRQEDQLQD YRKNNTYLVKRLEYESFQCGQQMKELRAOH
			1		<u> </u>	EENIKKLADQFLEEQKQETQKIQSNDGKELDI
			[ļ	NNQVVPKNIPKVAENVADKNEEPSSNHIPHG
88	1438	A	1218	1	534	PEFGTTISCGYLMATDVSRRPSVHKAVEIEOE
"		**	12.0	-		RVKSAGAWIIHPYSDFRFYWDLIMLLIMVGN
					Į.	LIVLPVGITFFKEENSPPWIVFNVLSDTFFLLD
			·		(LVLNFRTGIVVEEGAEILLAPRAIRTRYLRTW
·						FLVDLISSIPVDYIFLVVELEPRLDAEVYKTAR
•					}	ALRIVRFTKILSLLRL
89	1439	A	1223	1	743	MGFDEVFMINLRRRQDRRERMLRALQAQEIE
						CRLVEAVDGKVGMLTRSNAAPGRHLAMLET
						LVVVAPRFVDADNLILNPDTLSLLIAENKTVV
						APMLDSRAAYSNFWCGMTSQGYYKRTPAYI
						PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL
						\AFYPPHPDYTWSFDDIIVFAFSCKQ\AEVQMY
						VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ
						LEVMVPSSPSSAQSMAVVSADHIGLVISYL
90	1440	Α	1227	2	349	NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF
						GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/
]						YDRY/WKVVRHL/WDSWMTGI/SFTRVYLLG
<u> </u>	144		10.45		1000	LGARLVWFGKLILAKGGHGGISWL
91	1441	A	1245	3	1937	LGSSDVRAPQRSELGAESPSRMVASQAYNLT
						SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD
					ĺ	VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF
						QTQAGAQIPIERLASERAITVKVPNNSDWAAR
}						LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR
						PNEHNCSASRRIRPESLOGADHRPYTFFISPGT
1 1						RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC
	,	ļ				QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR
		i				HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML
j		.			.	TCAVCLVTYMVMAAILHKLDQLDASRGRAIP
1 1						FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM
		1			İ	LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP
		1		Ì		HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD
'		1			ļ	LQTARSTFFLVNDWLSVETEANGGLVEKEVL
.		Į			i	AASKASFRVPTPS\AALLRFRRLLVAELQRGF
]				FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL
		ſ			1	FLGANAVWYGAVGDSAYSTGRVSRLNPLSV
1	1	1	İ		[DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV
					l	GWGWGPGSTGNGAWASAPCPEPPLSSAAAR
					<u> </u>	GKGVHQRLLGKGQHT
92	1442	Α	1246	5	562	VFDEENILNELNDPLREEIVNFNCRKLVATMP
	ļ	j				LFANADPNFVTAMLSKLRFEVFQPGDYIREG
	1	.]	ļ			AVGKKMYFIQHGVAGVITKSSKEMKLTDGS
}	ĺ	1		l		YFGEICLLTKGRRTASVRADTYCRLYSLSVD
l	Ì	ļ				NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN
	1445					SILLQKFQKDLNTGVFNNQENEILKQIVKH
93	1443	A	1249	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL
1	1	ł	ł	}	ļ	PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN
1	1		1		· }	KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT
						ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

NO. of No. of not could	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Deptide Condition Contragnosis			1				
Sociation Soci]		nucleotide	1	
Sequence 994996 corresponding to the first of the side of the side of the side of peptide residue of peptide sequence Co-Gittamine, R-Anginie, S-Serine, T-Threonine, V-Vuine, W-Tryptophan, Y-Tryonine, V-Vuine, W-Tryptophan, Y-Tryonine, V-Vuine, W-Tryptophan, Y-Tryonine, V-Vuine, W-Tryptophan, Y-Tryonine, V-Vuine, W-Tryptophan, Y-Tryonine, V-Vuine, W-Tryptophan, Y-Tryonine, V-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryonine, W-Vuine, W-Tryptophan, Y-Tryptophan, W-Trypto			[
1440 1261 3 385 1711	1		ĺ				
### amino acid residude of sequence peptide residude of sequence peptide s		1	ł	1			
Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide muleotide felicion, Peposible muleotide filestrion Peptide sequence Peptide		1	i	**			
peptide							
	ł		1			sequence	
1444 A 1261 3 385	1	1	1				
				<u> </u>	sequence		
	1		l	[•		
1444 A 1261 3 385	l i	İ	i	1			
1444 A 1261 3 385	1	1		i		ŀ	1
1445 A 1282 2 250 350	-	1111					
95 1445 A 1282 2 550 GFRDNRGGEDPRETEVEHFGLVARFA CR GFRDNRGGEDPRETEVEHFGLWRFTFELVARFA VAPDFLKFFKNALNLDLMSIVPFYITLVVNL VVESTFILANLGRVAQVLRLMRIFRLKLARI STGRSLGATLKYSKEVGLLLLVLSVGISIPS VVAYTIEKEENEGLATIPACWWATVSMTT VQGDVVPGTTAGKLTASACLAA 96 1446 A 1294 1 1456 QLLPPSINENAGLLVGRCLCSAALRPVGDLIT SGQVAVRNAPQAGSAKAGRQDNFETIQ YFKKFFDANCNEKDYNPVAAGQGGFTEVAP SIVAPVLNKFNQCPEGYTCVAGNENNYGYT SEDTFSWAFLSLFRLMTQDYWENLVOLILAA AETTYMERLVA'LVLLGSLLVLLAVVAMA VEEQNOATLERAECKAEFQOMLEQLKXQQ EAAQQAATAASEHSREPSAAGRLSDSSSAS KLSSKSAKERRIRRKKRKQKQSGGEEKDEL EPQKSSESDSRRKGFRSIEGNRLTYERKYSS PHQSLLSIRGSLFSFRNNSTSIFSRGRAKDV GSENDFADDEHSTFEDNESRRDSLFPYRHGE RNNSLSJTSSSSRMLAVPFANGKMHSTVDC NGVSLVGGPSVTISVGQLLPEVILGVATD DINGTTTETEMRKRRSSSFITVSMOFLEDPSQR QRAMSLASILTITVE 97 1447 A 1295 2 2057 IQTQLPTKSSQOLREGGNCVRCKMGMNFIAE EVLLKYRITTYNNNKGPNILYTEKAPVHFMI NRYLSYGGGRKAFRPLVDVLVJELFASSKPV CTSPVDDIDASSPPSGSPSGTLPSTTGQGALS SELPSTSSVAALSSSSVALSFASSVALSFASSKVALVERHAV LVHBQQANAGHTWATIEDHESRWMKYNDJ AVTKSSWEELVRDSFGGYRNASAYCLMYIN DEAGGLEGDLIKTGQPLVGGTLPPDLRDFV EEDNORFEKELEEWDAQLAQKALQEKLLAS QKLRSSETSVTTAQAAGDPKYLEPGSRSDFSK HLKEETIQITKASHEHEKSSFTVLQSAIKLE YARLVKLAQEDIPPETDYRLILAVVYFIQNA QKKLEIKSTEVTLAQAAGDPKYLEPGSRSDFSK HLKEETIQITKASHEHEKSSFTVLQSAIKLE YARLVKLAQEDIPPETDYRLILAVVYFIQNA QKKLEIKSTEVTLAQAAGDPKYLQPGRSDFSK HLKEETIQITKASHEHEKSSFTVLQSAIKLE YARLVKLAQEDIPPETDYRLILAVDAKELL SKGLYRGHDEELISTYRECLILANGQAAEL SKGLENIPPEGVILESTEVPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLLDCSMEKSFHEPPLLQEKLT DPLIKLDCSMEKSFHEPPLLQEKLT DPLIKLDCSMEKSFHEPPLLQEKLT DPLIKLDCSMEKSFHEPPLLQEKLT SGGEDEVNINGLINDRETVPFILLQHPED LANTCKQGSQRVQOFEDALQKLREMBAHGRV WSCDLLQVERGWQLQLDTKEELDSVBLD SIQAMNVALNTCSYNSILS COMPANALNTCSYNSILS	94	1444	A	1261	3	385	
DPNHFYAMDETHHKCLVLRIQEAFGNKLCPA	1		ł				, , , , , , , , , , , , , , , , , , , ,
95 1445 A 1282 2 550 GPRDNPGIEDPRFEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPILANLIGRVAQVLRIMRIPRILKARI STGIRSLGATILKYSYKEVGLLILVLSVGISIPS VVAYTIEKEENGLATIPACWWATVSMTT VQGDVVPOTTAGKLTASACKPOMPTEIQ 96 1446 A 1294 1 1456 QLLPPSINENAGLLVGRCLCSAALRPVGDLT SGQVAVRNAPQAGSALAGKPOMPTEIQ YFKKFFDANCNEKDYNPVAAGQGGTEVAP SIVAPVLMKPNQCPEGYICVKAGRNPNYGYT SPDTFSWAFLSLFRIMTQDYWENLYQLTRA AETTYMIFLVALVILLGSLYLVTLILLAVVAMA YEEQNQATLEBAEGKAEAFQOMLEQLKXQQ EAAQQAATATASEHISRFPSAAGRLSDSSSEAS KLSSSSAKERNRKKKRKQSGGEEKDEI EFQKSSESDSIRRKGFRSIEGNRLTYERRYSS PHOSLISIRGSLFPFRNSETSRGRKDV GESNDFADDEHSTFEDNESRRDSLFVPRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVQQLLPEVIIDKPATD DNGTTTETEMRKRSSSHVSMDFLEDPSQR QRAMSLASLITATIVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFILE EVALLKYRITEYNNKGPNALVJEKAFVFFM NRYLSYGGPKRPPLVDVLQYALEFASSRV CTSPYDDDASSPPSGSPSGTSTTQGQALS SELPSTSPSSVAAISSRSVHKVFTGSREPDLP MHPAPRHITEELSVLESCLHRWTEIENDTR DLQESISRHRTIELMYSDKSANQVPYRLHAV LVHEQQANAGHYWAYIPDHRESRWMXYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYN DKAQFLIQEDLIKTGQPLVGETLPPDLENDTR DLQESISRHRTIELMYSDKSANQVPYRLHAV LVHEQQANAGHYWAYIPDHRESRWMXYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYN DKAQFLIQEDLIKTGGPLVGETLPPDLENDTPV EEDNQFFKELEEWDAQLAQKALQEKLLAS QKLRSSETSVTTAQAAGDPKYLEOPRSDFSK HKEETIQITKASHHEPUKSPSTVLQSAKLE VARLVKLAQEDTPPTTYPLHLIVVOYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKCVA QAKLEMIKPEVNLEEYEEWHQDYKKFRETT MYLIGLENFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLIKLNEQAAELF ESGEDREVNNCLIINMERIVPFIPLLLLYDEMEE KDILAVEDMRNWCSVLGGGORPHLQEKLT DFLYKLLLOCSMERSFHEPPKLPSVSTHELCER FARIMLSLSRTPADOR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHTVEH LMTCKQGGQRVQGFEDALQKLFEMDAHGRV WSQDLLLQVRIGWUQLLDLETKELDSYRLD SIQAMNVALNTCSYNSILS SIQAMNVALNTCSYNSILS SIQAMNVALNTCSYNSILS SIQAMNVALNTCSYNSILS SIQAMNVALNTCSYNSILS		1	ļ	}		j	
95	1 .		!				
VAPPELKFEKNALNI.IDLMSNYFFYTTI.VVNI. VVESTPTLANLGRVAQVVRLMRIFRILKLARH STGLRSLGATIKYSYKEVGLLLYJLSVGISIFS VVAYTIEKEENEGLATIPACW WAAVTSMTT GYGDVVPGTTACKLTASACILA 96 1446 A 1294 1 1456 QLLPFSNERNAGILVGRCLCSAALRFVGDLT SKOQVAVRNAPQAGSAKAGKGKPODNFEFIQ YFKKFDANCNERDYNPVAAGQGGETEVAP SIVAPVIANKPNOCPEGPICYACKAGRNPNYGYT SPDTFSWAFLSLFRLMTQDYWENLYQLTLRA AEITYMIFLV/LVILLGSLYJLVTLLLAVVANA YEEQNQATLEAEQKE AERQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRILSDSSEAS KLSSKSAKERSRRKRKKRKGKEOSGGEEKDEE EPOKSESEDSIRRKGFRSIEGRRLTYERRYSS PHOSLLSIRGSLFSPRRNSKIFSFRGRAKDV GSENDFADDEHSTFEDNESRRDLFYPRHGE RRNSNLSOJTSRSSMLAVFRNGKMSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR RRNSNLSOJTSRSSMLAVFNOKMMSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETTEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNIVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLKGGNCVRCKMQMNFIAE EVLLKYRITYNNNKGPMMLYEIKAFVHFM NRYLSYGSGPKRFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGISPSMMLYBEIKAFVHFM NRYLSYGSGPKRFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGISPSMMLYBEIKAFVHFM NRYLSYGSGPKRFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGISPSMMLYBEIKAFVHFM NRYLSYGSGPKRFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGISPSMMLYBEIKAFVHFM NRYLSYGSGPKRFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGISPSNMLYBEITEINDTR DLQESISRHRITIELMYSDKSMQVFYRLHAV LVHEGQANAGHYWAYIFDIRESRWMKYNDI AVYKSSWEGLEVDNSGGFCRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRHRITIELMYSDKSMQVFYRLHAV LVHEGQANAGHYWAYIFDIRESRWMKYNDI AVYKSSWEGLEVDNSGGFCPSGPSSK HKKEETIQLIFKASHEHENDRSAVCLMYIN DKAQFLIQEDLIKTGQFLVGIETLIPDLRDV EEDONQFFEKLEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEGPSGPSDFSK HKKEETIQLIFKASHEHENDRSPETVLQSAKILE YARLVKLAQEDTPEPTDYRLHHVVVYFIQMQ APKKIIEKTLLEGFGDRNLSFPERCHNIMKVA QAKLEMIRPEEVNLEEVEEWHQDYRKFREIT MYLIGLENRORSSYDDSLLFILCAYQNNKELL SKGLYRGHDEELISHYNRECLLKLINEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLUDEMEE KDDLAVEDMRNKVSTLQGEMPHLQEKLT DFLPKLLDCSMEIKSFHEEPRLPSYSTHELCER FARIM_SLSTRADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LLMTCKQGSQRVQOPEDALQKLFENDAHGRV WSQDLLQVORGWCLQFEVPPDLLRTTALGVHFE							
96 1446 A 1294 1 1456 SICHERCH STOCKSCHATT STANLERY STOCKSCHATTLY SYSTEM STOCKSCHATTLY SYSTEM STOCKSCHATT STANCE S	95	1445	Α	1282	2	550	GPRDNPG\EDPRFEIVEHFGIAWFTFELVARFA
STGLRSLGATLKYSYKÉRGILLLYSVGISIPS VVAYTIEKERNEGLATIPACWWWATVSMITT VGYGDVYPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QILPPSNRENAGILUGRCICSAALRIPVGDITT SSGQVAVRNAPQAGSAKAKGKFODNFETIQ YFKKFPDANCNERDYNPVAAGGGGETEVAP SIVAPVINKPNQCPEGYTUAKGRIPNYGYT SPDTTSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMFIA/VLULIGSVAKARINPNYGYT SPDTTSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMFIA/VLULIGSVAKARINPNYGYT SPDTTSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMFIA/VLULIGSVAKARINPNYGYT SPDTTSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMFIA/VLULIGSVAKARINPNYGYT SPDTTSWAFLSLFRLMTQDYWENLYQLTLRA YEEQNQATATASEHSREPSAAGRLSDSSSAA KLSSKSAKERRNRKKKRÇKEQSGGEEKDEL EPÇKSESEDSIRRKGFRSIEGNRLTYERYSS PHQSLSIRRGSLFSPRRNSINFSFRGRAKDV GSENDFADDEHSTFEDNESRRDLFYPRHGE RRNSINSOTSRSSMLANFIAFSFRAKOV GSENDFADDEHSTFEDNESRRDLFYPRHGE RRNSINSOTSRSSMLANFIAGKHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSFHVSMGMISTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSPTLYSTNGKMMSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD NRTLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSOTLPSTTGQGGALS SELPSTSPSSVAAISSRSVHKPFTGGRIPDLP MHPAPRHTEEBISVLESCLHRWRTEIENDTR DLQESISRHRTTELMYSDKSMQVPPVLHAV LVHQQANAGHYWAYIFDIRESRWMSYNDI AVTKSSWEGLINGVRGSPDDLFSTFPLDRFV EEDNQEFKEKLEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPFYLEGDFRNMKVA QAKLEMIKPEEVILESPECHNIMKVA QAKLEMIKPEVNLEEYEEWHQDYRKFRETT MYLLIGLENGRAELISH WRSCLLKLINEQAAELF ESGEDREVNNGLIMNEFTVPFLPLLLVDEMEE KDILAVEDMRNKVCSYLQAKALIL SKGLYKGHDEELISHYRRECLLKLINEQAAELF ESGEBREVNNGLIMNEFTVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGEMPHILQEKLT DFLPKLLDCSMEKSTHEPPKLPSYSTHELCER FARIM_SLSSTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFERDAHGRV WSQDLLQVORDWLQLLDEITKEELDSYKLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAFQAPPCPVPPDLLRTTALGYHFE]			}			VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL
96 1446 A 1294 1 1456 QLLPPSNRENAGILVGRCLCSAALRPVGDLIT SCQVAVRNAPQAGSAKAGKGKPONFERIQ YFRKFFDANKONERDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRINNYGYT SPOTTSWARLSIFRLIMTQDYWENLYQLTLRA AETTYMIFALVLVILIGSLYLVTILIAAVVAMA YEEQNQATLEEABQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREFSAAGRLSDSSSEAS KLSSKSAKERNRKKKRQKEQSGEKDED EPQKSSESDSSTRKGFRISPSAGRLSDSSSEAS KLSSKSAKERNRKKKRQKEQSGEKDED DEGNGSSESDSFRKGFRISPSAGRLSDSSSEAS KLSSKSAKERNRKKKRQKEQSGEKDED EPQKSSESDSFRKGFRISPSAGRLSDSSSEAS KLSSKSAKERNRKKKRQKEQSGEKDED DEGNGSSESDFADDEHSTEPDLSRDSLYPYRHGE RRISNLSQTSRSSSEMAAVFFANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSFNSMFLEDFSQR QRAMSIASILTNIVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQTRKGGNCVRCKMQMNFIAE EVILKVRITFYNNKGPNMLYIEIKAFVHFM NRYLSVGSGPKRPPLVDVLQVALEFASSKPV CTSPVDDIDASSPSGSPSQTLPSTTEQQALS SELPSTSPSSVAAISSRSVHKPFTGSRPPDLP MHPAPRHITEEEBSVLESCHHWRTEIENDTR DLQESISRHRTIELMYSDKSMQVYRLHAV LVHEGQANAGHYWAYIDFIRESRYMKYNDI AVTKSSWELVRDSFGGYRNASAYCHMYIN DKAQFLIQBOLIKTGQPLAGAGDPRYLEGPSRSDFSK HLKEETIQIITKASHEEBUSLESCHIPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPRYLEGPSRDFSK HLKEETIQIITKASHEEHBUSLFPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPRYLEGPSRDFSK HLKEETIQIITKASHEEBUSLESCHIPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPRYLEGPSRDFSK HLKEETIQIITKASHEEHBUSLEYEEWHODYKKRETT MYYLIGGLENFORSYDSLILLEGYGDRNISPETIVLQSAIKLE YARLVKLAQEDTPFETDYRLHHVVYFIQNQ APKKIIEKTILEGFGDRNISPDERCHNIMKVA QAKLEMIRPEFVPPTLLAYONKELL SKGLYRGHDELISHYRRECLLKLNEQAARLF ESGEDRENNINGLILMFEFTPFLILLYDEMEE KDILAVEDMRRIKVSKGLYGHDELISHYRRECLLKLNEQAALEF ESGEDRENNINGLILMFEFTPFLILLYDEMEE KDILAVEDMRRIKVSKGLYGHDELISHYRRECLLKLNEQAALEF ESGEDRENNINGLILMFEFTPFLILLYDEMEE KDILAVEDMRRIKVSKGLGARPHLQEKLT DFLPRILLDCSMEKSFHEFPFRLPSYSTHELCER FARIMLSLSKTPLADGR							VVESTPTLANLGRVAQVLRLMRIFRILKLARH
96 1446 A 1294 1 1456 QLLPFSNRENAGILVGRCLSAALREVGDLIT SSQVAVRNAPQAGSAKAGKGFQNNFBFQ YFKKFFDANCNEKDYNPVAAGQGETEVAP SIVAPVLNKPNQCPBGYICVKAGRNPNYGYT SFDIFSWAFLSIFILMTQDYWENLYQUTLRA AETTYMIRJ VAVILIGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGKSISDSSSEAS KLSSKSAKERNRRKKKKKQKEQSGGEEKDED EFQKSESEDSIRRKGFRSISEONRITYEKRYSS PHOSLISIRGSLFSPRANSTSISFSFRGARADV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFANKGMSHSTVOC NGVVSLVGGPSVFTISPVGQLLPEVIIDKPAATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASLITNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMMFLAE EVLLKYRITTYANNKGPNMLYJEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPFSGSISGQTLFSTTEQQGALS SELPTSTSSSVAAISSRSVHKPFTGSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLOESISRIHRTTELMYSKMJQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYNGETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESSFSVTTAQAAGPYLLEOPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHRHVVYFQNQ APKKITEKTLEGFGBNASFDERCHNINKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGGLENFORESYIDSPRSCHSPKE LSGGDREVNNGLIINNEFIVPFLPLLLVDEMEE KDILAVEDMRNRVCSYLGQEMEPHLQEKLT DFLPKLLDCSMERSFHEPPKLPSYSTHELCER FARIMLSISRTPADGE 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRWEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SQAMMVALNITCSYNSILS 99 1449 A 1306 3 1660 CGYYCHTCAPQAPPCPPPPDLLERTALGVHFE			ļ	[STGLRSLGATLKYSYKEVGLLLLYLSVGISIFS
96 1446 A 1294 1 1456 QLLPFSNRENAGILVGRCLSAALREVGDLIT SSQVAVRNAPQAGSAKAGKGFQNNFBFQ YFKKFFDANCNEKDYNPVAAGQGETEVAP SIVAPVLNKPNQCPBGYICVKAGRNPNYGYT SFDIFSWAFLSIFILMTQDYWENLYQUTLRA AETTYMIRJ VAVILIGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGKSISDSSSEAS KLSSKSAKERNRRKKKKKQKEQSGGEEKDED EFQKSESEDSIRRKGFRSISEONRITYEKRYSS PHOSLISIRGSLFSPRANSTSISFSFRGARADV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFANKGMSHSTVOC NGVVSLVGGPSVFTISPVGQLLPEVIIDKPAATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASLITNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMMFLAE EVLLKYRITTYANNKGPNMLYJEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPFSGSISGQTLFSTTEQQGALS SELPTSTSSSVAAISSRSVHKPFTGSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLOESISRIHRTTELMYSKMJQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYNGETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESSFSVTTAQAAGPYLLEOPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHRHVVYFQNQ APKKITEKTLEGFGBNASFDERCHNINKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGGLENFORESYIDSPRSCHSPKE LSGGDREVNNGLIINNEFIVPFLPLLLVDEMEE KDILAVEDMRNRVCSYLGQEMEPHLQEKLT DFLPKLLDCSMERSFHEPPKLPSYSTHELCER FARIMLSISRTPADGE 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRWEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SQAMMVALNITCSYNSILS 99 1449 A 1306 3 1660 CGYYCHTCAPQAPPCPPPPDLLERTALGVHFE	}	}	ł			,	
SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMFULVVILLGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERNRRKKKQKEQSGGEKDED EFQKSSSEDSIRRKGFRSIEGINTLYEKRYSS PHQSLLSIRGSLFSPRRNSITSLFSFRGRAKDV GSENDFADDEHSITEDNESRNDSLFVPRRHGE RRNSLSOTRSSSRMLSYPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASLITNTVE 97 1447 A 1295 2 2057 IQTQLFTKSSQQLRKGGNCVKCMQMMFIAE EVLLKYRITTYNNNKGMPULYELKAYHFMI NRYLSYGSGPKRPPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLFSTTEQQGALS SELPTSTSSSVAAISSRSVHKPFTQSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTTELMYSKMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYN DKAQFLIQEDLIKTGQPLVGIETLPPDLEDFV EEDNQRFEKELEEWDAQLAGKALQEKLLAS QKLRESETSVTTAQAAGRYLLEOPSSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETTYRLHHVVVYFIQNQ APKKITEKTLLEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVWLEEYEEWHOPYNKPEIT TMYLIGGLBFORESVIBSLLFILICAYONNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIINNEFIVPPPLLLVDEMEE KDILAVPDMRNRWCSYLQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSISTTPADGE 98 1448 A 1304 118 453 SGPSSRATYLFIRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQPPDALQKLTEPDMAHGRV WSQDLILQVRQGPULQLDIETKEELDSYRLD SIQAMMVALNITCSYNSILS 99 1449 A 1306 3 1660 CGYYCHTTCAPQAPPCPYPPDLLETTALGYHFE			1		!		VGYGDVVPGTTAGKLTASACILA
SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMFULVVILLGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERNRRKKKQKEQSGGEKDED EFQKSSSEDSIRRKGFRSIEGINTLYEKRYSS PHQSLLSIRGSLFSPRRNSITSLFSFRGRAKDV GSENDFADDEHSITEDNESRNDSLFVPRRHGE RRNSLSOTRSSSRMLSYPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASLITNTVE 97 1447 A 1295 2 2057 IQTQLFTKSSQQLRKGGNCVKCMQMMFIAE EVLLKYRITTYNNNKGMPULYELKAYHFMI NRYLSYGSGPKRPPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLFSTTEQQGALS SELPTSTSSSVAAISSRSVHKPFTQSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTTELMYSKMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYN DKAQFLIQEDLIKTGQPLVGIETLPPDLEDFV EEDNQRFEKELEEWDAQLAGKALQEKLLAS QKLRESETSVTTAQAAGRYLLEOPSSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETTYRLHHVVVYFIQNQ APKKITEKTLLEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVWLEEYEEWHOPYNKPEIT TMYLIGGLBFORESVIBSLLFILICAYONNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIINNEFIVPPPLLLVDEMEE KDILAVPDMRNRWCSYLQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSISTTPADGE 98 1448 A 1304 118 453 SGPSSRATYLFIRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQPPDALQKLTEPDMAHGRV WSQDLILQVRQGPULQLDIETKEELDSYRLD SIQAMMVALNITCSYNSILS 99 1449 A 1306 3 1660 CGYYCHTTCAPQAPPCPYPPDLLETTALGYHFE	96	1446	A	1294	1	1456	
YFKKFFDANCNEKDYNPVAAGOGOETEVAP SIVAPVLINKPNOCPSTOVAKAGRINPNOCYT SFDTFSWAFLSIFRIMTODYWENLYQLTIRA AEITYMIFILVILVILIGSLYLVILIILAVVAMA YEEQNQATLEAEAGKEAFFQOMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRINRKKRKQKEQSGGEKDED EPQKSESEDSTRRKKGRKFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRINSRTSLFSTRGRAKDV GSENDFADDEHSTFEDNESRDSLFYPRHGE RRINSILSOTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATD DNGTITTETEMRKRSSSHVSMDFLEDPSQR QRAMSIASILTNIVE 97 1447 A 1295 2 2057 IQTQLPYKSSQQLKGGNCVRCKMQMMFIAE EVILKYRITFYNNKGPNMLYIEIKAFVHPMI NRYLSYGSGPKFFLVDVLQVALEFASSKPV CTSFVDDDASSPPSSGPSGTLFSTTEQQGALS SELFSTSPSSVAAISSRSVHKFTQSRDPDLP MIPAPRHTTEELS VLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVARDSGFGYNNASQYCLMYIN DKAQFLQEDDLKTGQPLVGGTLPPDLRDFV EEDNQRFEKELEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKESTIQHITKASHEHEDKSPTVLQSAIKLE YARLVKLAQEDIPPETDVRLHHVVVYFIQNQ APKKHEKTLLEQFGRONLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENPQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMDFFIPPIPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFIPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 GGPSSRAYLHRKEYSQNLTSSPTLLQHRVEH LMTCKQGSORVQOPEDALQKLFREMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHFE	1		ł				
SIVAPVLINKPNOCEGYICVKAĞIRİPINGYT SFDTFSWAFLSIFRIMTODYWENLYQLTILRA ABITYMIRI.VILUĞSLYILVTILLAVVAMA YEQNQATLEAAGQKEAEFQQMLEQLKKQQ EAAQQAATATASERPSAAGRI SDSSSEAS KLSSKSAKERRIRKKRKQKEQSGGEKDED EFQKSESEDSIRRKGFRSIEGIRILTYEKRYSS PHOSILSIRGSLIFSPRINSITSLIFSTRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRINSILSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVTFSPVGQLIPEVIIDKPATID DINGTTTETEMRKRRSSSFHVSMDFLEDFSQR QRAMSIASILTITVE 97 1447 A 1295 2 2057 İQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLİKYRITEYNINKGPINMLYIEIKAFVHFMI NYLSYGSGPKRFPLVDVLQYALLEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSVAAISRSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEQQANAGHTYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGETLPPDLRDFV EEDNQRFFKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPFDDVRLHHVVVYFIQNQ APKKIEKTILEQFGORNLSFDERCHNIMKVA QKALEMIKPEEVULEYEEWHQDYKFRETT MYLIĞLENFQRESYIDSLIFLICAYQNINKELL SKGLYRĞIDEELISHYRRECLIKLINEQAAELIF ESGEDREVNINGLIMNEFIYPFLPLILYDEMEE KDILAVEDMRINEWCSYLĞGÇMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 GGFSSRAIYLHRKEYSQNLTSEPTILQHRVEH LMTCKQGSQRVQQPEDALQKLFRMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALITCSYNSILS 199 1449 A 1306 3 1660 CGYFCHTTCARQAPPCPVPPDLLRTALGVHFE	}]			
SFDTFSWAFLSLFRLMTODYWENLYQLTIRA ARTTYMFILVI/VILLGSLYLVTILLAV/VAMA YEEQNQATLEAAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRRKKKQKEQSGEBKDED EFQKSESEDSIRRKGFRFSIEGINRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSSMLAVFPANGKMHSTVDC NGVVSLVGGFSVFTVGQLLFEVIDDKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 1447 A 1295 2 2057 IQTQLPTKSSQQLFKGGNCVRCKMQMNFIAE EVILKYRITFYNNKGPNMLYEIKAFVHPMI NYLSYGSGPKFRPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTITEQQGALS SELPSTSPSSVAAISSRSVHKFTQSRPPDLP MIPAPRHITEELSJCLSCLHRWRTEENDTR DLQESISRHRTIELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVABDSFGYNNASAYCLMYIN DKAQFLQEDDLIKTGQPLVGGTLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HIKEETTIQIITKASHEHEDKSPTVLQSAIKLE YARLVKLAQEDTPPFTDVRLHHVVVYTQNQ APKKIEKTLLEQFGGRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEYFEWHQDYRCFETT MYLIGLENFQRESYIDSLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNFFTVPFIPLILUTDEMEE KDILAVEDMRNRWCSYLGGEMEPHLQEKLT DFIPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 GGFSSRAYLHRKEYSGNLTSFTTLQHRVEH LMTCKQGSGRVQOPEDALQKLEFMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SQAMNVALATICSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHFE							
AETTYMIFIA-VILVILLGS.YL-VTLILA-VIVAMA YEEQNQATLEAEQQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSRSAKERRNRKKKRQKEQSGGEEKDED EFQKSSEDSIRKGFRSTEIGENRLTYEKRYSS EHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFYPRHGE RNNSILSQTRSSSRMLA-VFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATD DNGTITTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPYRSQQLRGGGNVCRCKMQMNFIAE EVILLKYRITFYNNKGPNMLYIEIKAFVHPMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDIDDASSPPSGISPSQLT-PSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTGSRIPPDLR MHPAPRHITEELS-VLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDSSMQLVPYKHAPV LVHEQQANAGHYWAYIFDIRESR WMKYNDI AAVTKSSWEEL-VRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQFSRSDFSK HLKEETIQITKASHBHEDKSPET-VLQSAKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNIS-DERCHNIMKVA QKKLESTISVTTAQAAGDPKYLEQFSRSDFSK HLKEETIQITKASHBHEDKSPET-VLQSAKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNIS-DERCHNIMKVA QKKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESVIDSLLFILCAYONNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFI-PLLLVDEMBE KDILA-VEDMRNRWCSYLOGEMEPHLQEKLT DFLPKLLDCSMERKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNT.SEPPILLQIRVEH LMTCQGOSRQVGPEDALQKLFEMDAHGRV WSQDLILQVEDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1						
YEEQNQATLEEAEQKEAEFQOMLEQLKKQQ EAAQQAATATASEHSREPSAAGRISDSSSEAS KLSSKSAKERRINRKKKRKQKEQSGGEEKDED EFQKSESEDSTRRKGFRFISIGNRLTVEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTIEDNESRRDSLFVPRRHGE RRNSILSQTSRSSRMLAVFANGKMHSTVDC NGVVSLVGGPSVPTSPVQQLLPEVIDDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRGGONCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHPM NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSSPOOTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHFARHTIEEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTQQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFBRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLSHFQRESYIDSLIFLICAYQNNKELL SKGLVRGHDEELISHYRRECLIKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLQGEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCXQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDWQUGPULQILDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCVPVPDLLTRALGVHPE	1 1						
EAAQQAATATASEHSREPSAAGRISDSSSEAS KLSSKSAKERRNRKKKQKEQSGGEKDED EPQKSSEDSIRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSTSTISFSFGRAKDV GSENDFADDEHSTFEDNESRFDSLFYPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVFTSPVGQLLPEVIDKPATID DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVILKYRITTYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSV AAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLQEDIKTGQPLVGIFLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITIKASHEHEDKSPETVLQSAKILE YARLVKLAQEDTPPETDYRLHHVVVYFTQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEVNLEYEEWHQDYKKPRETT MYLIGLENFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNCLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SCPSSRATYLHRKEYSONLTSEPTLLQHRVEH LMTCKQGSGVQFDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAFQAPPPCPVPPDELRTALGVHPE	1						
KLSSKSAKERNRRKKRKQKEQSGGEKDED EPQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDXPATD DNGTITETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTITEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITESLSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIPDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTAQAAGDPKYLEQPSRSDPSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLAGEFGPRNLSFDERCHIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRSSYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLIKLNEQAAELF ESGEDREVNNGLIMNEFIVPFI,PLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARINLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTICSYNSILS 99 1449 A 1306 3 1660 CGYFCHTICAPQAPPCPVPPDLLRTALGVHPE							
EPÇKSESEDSIRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATID DNGTITETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSSPSQTLPSTITEQQGALS SELPSTSPSSVAISSRSVHKPFTQSRPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIPDHESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLQEDDLIKTGQPLVGIETLPPDLRDFV EEDNQRFSKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAKKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEFYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNKCSYLGQEMEPHLQEKLT DFLPRLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTICAPQAPPCPVPPDLLRTALGVHPE	1 1						
PHÓSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSOTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSLASELTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLFSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELLSYLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDHIKTQGPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAGDPKYLEQPSRSDFSK HIKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDITPFTDYRLHHVVVYFIQNQ APKLIEMILEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLEINFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFPLLLLVDEMBE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARINLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTILQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
GSENDFADDEHSTEDNESRRDSLFYPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFFLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSGTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHTIEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAVCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFFKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEWHQDYRKFRETT MYLIGLENRQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIINNEFIVPFLPLLLVDEMEE KDDLAVEDDMRNRWCSYLGGEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGPSPVTSPVGQLIPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLKKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKITLLEQFGDRNLSFDERCHNIMKVA QAKLEMFEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLVDEMEE KDILAVEOMRNRWCSYLGQEMEPHLQEKLT DFLFKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAYLHRKEYSQNLTSEPTILQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTCAPQAPPCPVPPDLLRTALGVHPE				-			
NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNIKGPNMLYIEIKAFVHFMI NRYLSYGSGPKFPLVDVLQVALEFASSKPV CTSFVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETIPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQITKASHEHEDKSPFTTVLGSAKLE YARLVKLAQEDTPPETDYRLHHVVVYFJQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKPRETT MYNLIGLENFQRESYIDSLLFILCAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFTVPFIPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAVLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTCAPQAPPCPVPPDLLRTALGVHPE							
DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLQESRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QKALEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIINMSEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHITCAPQAPPCPVPPDLLRTALGVHPE							, ,
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE PVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAISSRSVHKPFTQSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRHRTTEELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFILPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1						
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHTIEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLQQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
EVLLK YRITFYNNNKGPNML YIEIKAFVHFM NRYLSYGSGPKRFPL VDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQITPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHR WRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFPHRESR WMK YNDI AVTKSSWEEL VRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPL VGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	07	1447		1000		0065	
NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSSVAAISSRSVIHKPFTQSRPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	97	1447	A	1295	2	2057	
CTSPVDDIDASSPPSGIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHIKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQOPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE						:	1
SELPSTSPSSVAAISSRSVIIKPFTQSRIPPDLP MHPAPRHITEELSVLESCHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLINEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1	•					
MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HI.KEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNINGLIIMNEFIVPFLLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPFKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	!!			1			
DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HI.KEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFIPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		,	İ				
LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							DLQESISRIHRTIELMYSDKSMIQVPYRLHAV
DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1 1						LVHEGQANAGHYWAYIFDHRESRWMKYNDI
EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			i		J		AVTKSSWEELVRDSFGGYRNASAYCLMYIN
EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					ļ		
QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKITLIEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	.			1	ł		
HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE)			J	J		
YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNINGLIIMNEFIVFFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	j l			[j		• • • • • • • • • • • • • • • • • • • •
APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	[-	i	[{	
QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					1		
MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE]]	j	J	<u> </u>	J	l	
SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPFKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					4	ĺ	
ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1 1	ł	- 1		1	ľ	
KDILAVEDMRNRWCSYLGQEMEPHLQEKLT	[1	į	1	j	1	ļ	`
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE) [•		
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE] [ĺ		ĺ		[
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE				ŀ		. 1	
LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	J	1440		1204	170	453	
WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	98	1448	Α	1304	112	403	SUPSKAIYLHKKEYSQNLTSEPTLLQHRVEH
99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			1	1		1	
99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		1	Ì	ļ		l	
The state of the s	L						
TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS	99	1449	A.	1306	3	1660	
						l	TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l		•		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			i I	peptide	į	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			1			DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF
1		ł				SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP
			ļ			TTCTVLLLAESEGERERWLQVLGELQRLLLD
			ĺ		ĺ	ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD
]			QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ
i i						QLTLSPSAGLLVVLCGRGPSVRLFALAELENI
						EVEVPKIPESRGCQVLAAGSILQARTPVLCVA
						VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ
1				1		SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL
						GAGLVPEELPPSRGGLGEALGAVELSLSEFLL
i i						LFTTAGIYVDGAGRKSRGHELLWPAAPMGW
				'		GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL
						KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE
						KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE EQOKQQRREMLKDPFVRSKLISPPTNFNHLV
						HVGPANGRPGARDKSP
100	1450	Α	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPOA
100	1470	^	1210	210	190	LNDSMINETARDAARVQVASTLSVLVGLFQV
			1			GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF
						VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL
						PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ
				1		OLPMPIPGELLTLIGATGISYGMGLKHRFEAG\
						PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK
1						IFALRHGYRVDSNQVWVMRDV
101	1451	Α	1353	220	445	DWPDLFTYPLIGSPKCFQSARPE\RMYRRTVR
						SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD
1		į				HGHPPARMSIFSR
102	1452	Α	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN
1			·			WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY
1						CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA
1						VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW
		. '				YNLILTA\LLCLPLVIVTLCYTTIIHTLTHGHAN
L						\DSCLKQKARRLTILLL
103	1453	A	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQV\C
					,	SFNEHGYHLFQAMRLAVEEINNSTALLPNITL
					!	GYQLYDVCSDSANVYATLRVLSLPGQHHIEL
						QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP
						FLVPMLLEQ
104	1454	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG
						FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP
						RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW
				İ		ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG
105						KSREERFCNENTPCPVPIF
105	1455	A	1379	2	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD
						IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF
						SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK
						CQSDWMERWVDDAFWSFLF\SLILIVIMFLW
100	1166		1202	- .	400	RPSA
106	1456	Α	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW
				ı		QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG
						CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL
						FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF
105			100/			QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN
100	1456		100=			YPALSLQSSWDHRHTWLIFAFL
108	1458	Α	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM
				ľ	1	VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT
				l		VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK
L		لــــــــــــــــــــــــــــــــــــــ				AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in No.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	umac	ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Juance	}	1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ļ	ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide	Seduction	
	İ		1	sequence	ì	nucleotide insertion
	 		 	sequence		FQQMLGQGIAGILPKLIGGYFDTDQRAAGLG
1	ļ					FTYNVGALGGALAPIIGALIAQRLDLGTALAS
1	ŀ		1	ļ	}	LSFSLTFVVILRNRRPGKSLVR
109	1459	A	1402	15	387	VLVALPDT\VTSETVVTEVLGHRVTLPCLYSS
***	1,	1	1 1 102	**	307	WSHNSNSMCWGKDQCPYSGCKEALIRTDGM
	ļ	j				RVTSRKSAKYRLQGTIPRGDVSLTILNPSESDS
		i i				GVYCCRIEVPGWFNDVKINVRLNLORASTT
110	1460	A	1421	3	350	HEDLSSLLTRGSGNOEREROLKKLISLRDWM
	1	1 **	1	~	1 330	LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF
ļ						FFHSPDALFSLLLLSCYFPSYCFFYYLFFSSSPL
ļ	1	1	1			CLLLASSPFPLFILLASL
111	1461	A	1426	2	344	FTSTMTKPFEKESEQPA*ATLAFGAQTSTTAD
		١	``~~`	-		QCALKPDLSYLNNSSSSSSTPATSAGGGIFGSS
)	J				TSSSNPPVATFVFGQSSDPVSSYGFVNTAESST
		İ				SDSLLFSQDSKLATTS
112	1462	A	1434	46	372	TTSWTTSCTRSCT*SGASSGPGWTPRTTWWR
	1	{	1	10	3/2	SRRSSORTCSRACSGAWSRTW*RSS*TSSSSC
ļ					ì	STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC
ł	ļ	!	i .		ļ	MSSSTTSSTTSTF
113	1463	A	1439	3	292	HEDIMTHYDRLVDE*ALNAGKQRYEKMISG
	1	ļ · •	1 137		-/-	MYLGEIVRNILIDFTKKGFLLRGOISEMLKTR
	l	ŀ			ļ	GIFLTFLLSNFLIVCVLLFYVSFYLFOSCINFVL
114	1464	A	1463	1	396	KQQAVPEPHSSTTTPQEQEQNWYGQDLLNLQ
	[- 100	-	550	QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP
	l	ł	1		!	ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPOS
	Î	ľ			İ	EDP*KNA*LKQMHAATTHWQQHQQHQVGC
	ŀ					QYHGIMQ
115	1465	A	1464	291	2	AGSYPSMVWSCHWGVTOKRRAL*VYSFEEG
						GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN
						MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN
			[NYCN
116	1466	A	1465	667	337	LPPQRPA*TDSYSTCNVSSGFLAGOSHNIHLO
	i	Ì	1			YWTKYQVWEWLQHFLDTNQLDANCIPFOEF
i	[DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ
						HLKWNGDSLFLCLSLPC
117	1467	A	1479	1	381	GTSGGPKRVLVTERFPWQNPLPVNRGQAQR
1						VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK
]]					QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN
						NPEEELASDPNNEESL*RPWALEDFEIGRPLG
						KGK
118	1468	A	1485	3	385	TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS
}		·				PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL
						QHPWIEGHTCLDNNIHQAASEPINNNFAESKR
)						NLAFLATGVVRHMRKLFMGANLEGPGPTVS
			L			Н
119	1469	Α	1486	1	398	GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL
						NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE
			1			KALTKFLKWVNWDLPQEAKQALELLGKWK
						PMDVKDSLELLSSHYTNPTVRRYAVARLRQA
						DDEDLLMYL
120	1470	Α	1497	3	999	MGESPAV*GYFVLAGMNSAGLSFGGGAGKY
					}	LAEWMVHGYPSENVWELDLKRFGALQSSRT
						FLRHRVMEVMPLMYDLKVPHWDFQTGRQL
				1	ļ .	RTSPLYDRLDAQGARWMEKHGFERPKYFVP
1	1			.	. 1	PDKDLLALEQSKTFYKPDWFDIVESEVKCCK
	·		ľ	.[EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS
					1	NDLDVPVGHIVHTGMLNEGGGYENDCSIARL
				1		NKRSFFMISPTDQQVHCWAWLKKHMPKDSN
Ll						LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ŀ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence	<u></u>	nucleotide insertion
1			į į			MTPDHFPSLFCKEMSVGYANGIRVMSMTHT
121	1471	A	1498		306	GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	14/1	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN
						WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE MLVAHTHIVEEHTGTHLQYVSWPDHSVPDD
						SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT
					1	WDPGTDTALGWSKQPSQSYTLFES*VGSGYII
						DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE
1 1						RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA
						AKHGHSPAVQVLLAQWQDINEMNEKQQTPL
	1454		1.5.5.			HVAADRG
124	1474	A	1555	1	745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL
1						HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP YVKFRLGHQKYKSKIMPKTLNPQWREQFDF
						HLYEERGGVIDITAWDKDAGKRDDFIGRCQV
						DLSALSREQTHKLELQLEEGEGHLVLLVTLT
						ASATVSISDLSVNSLEDOKEREEILKRYSPLRI
						FHNLKDVGFLQVKVIRAEGLMAADVTGKSD
						PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL
						*VALVWKKFQTQSLRLSDLHRKSHLWRGIVS
						ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY
						KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT
]						AWDKDAGKRDDFIGRCQVDLSALSREQTHK
						LELQLEEGEGHLVLLVTLTASATVSISDLSVN SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV
						KVIRAEGLMAADVTGKSDPFCVVELNNDRLL
1						THTVYKNLNPEWNKVFTL
125	1475	A	1556	57	509	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA
l i					,	CGGLDNICSIYNLKTREGNVRVSRELPGHTGY
	1					LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT
				•	•	TTFTGHSGDVMSLSLSPDMRTFVSGACDASS
126	1476		1592		170	KLWDIRDGMCRQSFTGHVSDINAVS
120	14/0	A	1392	3	178	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL
127	1477	A	1612	1	497	EMLPTCDLADQHNIKFHYAFALNR*ER TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS
""	14''	^ [1012		477	VLGAYISFGVPSSHLLTASVMSAPASLAAAKL
		ļ	.			FWPETEKPKITLKNAMKMESGDSGNLL*AAT
1	Ì	- 1	i	(QGASSSISLVANIAVNLIAFLALLSFMNSALA
		į				WVGNMFDYPQLSFELICSYIFMPFSFMMGVE
				l		WPDSFM
128	1478	Ā	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA
		}	}		ļ	EDEVDFRASSISEEVAVGSIAATLKMKQGPM
120	1470	1	1605		205	TQAINR
129	1479	A	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI
		ĺ			Ì	MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA
						LLCVWALSLVIYIGPLLGWRHPAPEDETICQI NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV
1 1	1	. 1	İ			AKTE
130	1480	A	1638	2	466	DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG
1 1						ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS
		ŀ	1	ļ		EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHP
}	}	į	ļ	Ì		KPSEASTTLPESHLKTPQMPPTTPSSSSFTKVT
			l			KDKDIK*LLFNLYSSVEILPEVLHLKT
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFRQMVE
		Ì	Ì			AIRYCHGCGVAHRDLKCENALLQGFNLKLTD
	1	ļ	ł	1	ļ	FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ
						GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TDIPKMLWQQQKGVSFPTHLSISADCQDLLK
120	1400		1656	1.50		RLLEPDMILRPSIEEVSWHPWLAST**KQWQV LSNKVGGESKPKKKK
132	1482	A	1656	150	48	LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV VDAQ
133	1483	A	1660	3	406	RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF PNFTP
134	1484	A	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP FFPAGAPPASSSSSSSSSSSPPTVSTAPPLIPPPGF PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS SSSSSSSSSSSSSSPRDRDRER*RTREREREDHS PTPSVFNSDEERYRYREYAERGYERHRASRE KEERHRERRHREKEETRHKSSRSNSRRRHESE EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE STEATPAE
135	1485	A	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL GYRYWAGIGVLQSCESALTHYRLVANHVAS DISLTGGSVVQRIRLPDEVENPGMNSGMLQE DLIQYYQFLAEKGDVQAQVGLGQLHLHGGR GV*QNHQRAFDYFNLAA
136	1486	Α	1678	525		ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS GSSSTASSLNFSAIMGSSSATASWVLSTASTPP CPSALPSSPAQES*SLAASSSAWPVAGISPSGA CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD SSSLSL
137	1487	A	1680			AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE DNAELNNQNFYLSKQLDEASGANDEIVQLRS EVDHLRREITEREMQLTSQKQTMEALKTTCT MLEEQVMDLEALNDELLEKERQWEAWRSVL GDEKSQFECRVRELQRMLDTEKQSRARADQ RITESRQVVELAVKEHKAEILALQQALKEQK LKAESLSDKLNDLEKKHAMLEMNARSLQQK LETERELKQRLLEEQAKLQQMDLQKNHIFR LTQGLQEALDRADLLKTERSDLEYQLENIQV LYSHEKVKMEGTISQQTKLIDFLQAKMDQPA KKKKVPLQYNELKLALEKEKARCAELEEALQ KTRIELRSAREEAAHRKATDHPHPSTPATARQ QIAMSAIVRSPEHQPSAMSLLAPPSSRRESST PEEFSRRLKERMHNIPHRFNVGLNMRATKC AVCLDTVHFGRQASKCLECQVMCHPKCSTC LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT KEPSSSLHLEGWMKVPRNNKRGQQGWDRK YTVLEGSKVLIYDNEAREAGQRPVEEFELCLP DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN ENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNK FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning mucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YGRRSRTDDLKWSRLPLAFAYREPYLFVTHF NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA ISSGAIYLASSYQDKLRVICCKGNLVKESGTE HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK
138	1488	Ā .	1686	2	526	GRPQGPAPGAGSPPESGPGLWAALGCSLVWV PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS CWTRGCQTTARTAAAAAAPGPAGRRPPGGA PQNGSCAASASQEAAAPPPMCPPGRRWAVAS PPETRCPAAPGTRCRRLEAA
139	1489	A	1693	3	376	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE FIETKALGCAWFSLCYYLVLYFESSHKVDFVF IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR RWAGIAKGVGTQKIIGRVHLGEQKALGL
140	1490	A	1704	3	376	ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	A	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ DKLELELVLKGSYEDTQTSFLGTASAFRFHY MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV PLRPLTIVKEVTMDVPAPNVRGLNWMG
142	1492	A	1769	1	, 406	NNPSTLPRGS*PMSPRTTMGRRRQRRREHKSS LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG LLQVGDRVLSINGIATEDGTMEEANQLLRDA ALAHKVV
143	1493	A	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL LENCMEMHCMDLPTDTSALVR
144	1494	A	1814	1	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF KCNGEWVSQNDHVTQEGLDEATGLRVREVH IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI
145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCIHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL
146	1496	A	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET
147	1497		1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q
148	1498	A	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS
149	1499	Α .	1880	611	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP VQVISGEEEARLIYQGVAHTTGGADQRLVVD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-		1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
4	peptide	1	USSN			I=Isoleucine, K=Lysine, L=Leucine,
cotide	seq-	Ī		location	corresponding	
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uenœ	ì	i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
I	ļ	!	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ŀ	1	ļ.		peptide		/=possible nucleotide deletion, \=possible
1	l	l	1	sequence		nucleotide insertion
					 	IGGASTELVTGTGAQTT*LFSLSMGCVTWLER
		1		ŀ	1	YFADRNLGQENFDAAQKAAREVLRPVADEL
	ļ	[]	1		RYHSWKEVRGASVTVQALQEIMMAQGMDE
ł	ł	ì	l	ł	ļ ·	RITMEIWPVD
1.60	1500	 	3804		750	
150	1500	Α	1894	2	750	GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH
1		l		•	1	LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD
1		l	Ì		İ	EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS
ĺ		ĺ			[PESSILDGMIRQLQQQQDQRMGADQDTIPRG
1		l				LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV
1	Ì	Ì	1	i .		EGVRQMHQNAPRSQIATERDLQAWKRRVVV
i		!				PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK
1	i .	ł		ł		TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	A	1900	141	785	GKTIOIOTTMONKYKTVOKOYKTIPKNKKA
					1	MEMQIKKQFQDTCKVQTKQYKALKNHQLEV
	1	l				TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI
1		İ	į l		1	NEMMASQALRLDEAQEAECQALRLQLQQEM
ł			1	i		
İ		İ		ļ	ł	ELLNAYQSKIKMQTEAQHERELQKLEQRVSL
1			1	Į		RRAHLEQKIEEELAALQKERSERIKNLLERQE
						REIETFDMESLRMGFGNLVTLDFPKEDYR
152	1502	Α	1915	2	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL
1			i			RQKIFKERALPDIENYMFENHDQLRQAATEC
j					[MCNMVLHKEVQERFLADGNDRLKLVVLLCG
1						EDDDKVQNAAAGALAMLTAAHKKLCLKMT
ĺ			1		ĺ	OVTT
153	1503	A	1921	1	237	AYOSLRLEYLQIPPVSRAYTTACVLTSAAVOL
'''		'`		•	1	ELITPFOLYFIPELIFKHFOIWRLITNFLFFVPFG
					·	FNFLLYMIFLYT
154	1504	A -	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR
154	1304	^	1926	-	334	
1			1		1	YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV
			[RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS
			<u> </u>			QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA
]						KSEALVLREKSTLERIHKHQEIETKEIYAQRQ
		i '			i '	LLLKDMDLLRGREAELKQRVEAFESYQLELK
]]			DDYIIRTYRLIEDDRINIQISGHWQESP
156	1506	À	1935	1	270	VTRKLPIFIVDAFTARAFRGSPAADCLLENEL
			1			DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ
						RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	A	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA
1.57	1507	^	1730	207	303	AYEYYDAGNHWCKDCNTICGTMFDFFTHMH
		l				
160	1600		1000	ļ. ——	400	NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	A	1939	1	423	TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA
·						NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED
[RLHCQTQACPPLSWPQRLDILLGTARAIQFLH
1						QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA
						RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR
	,					NGHLATVKLLVEEKADVLARGPLNQTALHL
						AAAHGHSEVVEELVSADVIDLFDEQGLSALH
						LAAQGRHAQTVETLLRHGAHINLQSLKFQGG
1 1						HGPAATLLR
160	1510	_	1002	2	417	
100	1210	A	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA
				İ		IYSEYCNNHPGACLELANLMKQGKYRHFFEA
j i						CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL
j í	İ					LKYTTQEHGDYSNIKAAYEAMKNVACLINER
<u> </u>						KRKLESIDKIA
161	1511	A	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE
<u> </u>						LWLETLQGQRHSHTGVKSTPGQSAAILMKLR
1 1						SSHNASKTLNANNMETLIECQSEGDIKEHPLL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}		}] .	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			· '	peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
}	<u> </u>		 	sequence	 	ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
				·	1	PASDFSGALETDLKASLFDQPLSIICGDSDTLP
Į						RPIODILTILCLKGPSTEGIFRRAANEKARKEL
1					ĺ	KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP
1			[RKLLSSDLFEEWMGALEMQDEEDRIEALK
162	1512	Ā	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI
		1	.,,,,		1	KFQGRWGTVCDDNFNIDHASVICROLECGSA
1 ,						VSFSGSSNFGEGSGPIWFDDLICNGNESALWN
ļ :						CKHQGWGKHNCDHAEDAGVICSSKD
163	1513	A	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPQPAATNGD
					107	LASRSNIAFMGTLVRCGKAKGVVIGTGENSE
1			1			FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAQLFGG
1		_	[1.5	SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI
						ENNWYFVVADSSKAGFTTIYKWERETGFYSH
1			'			QSFTR
165	1515	A	2013	2	403	EDPEELGHFYDYPMALFSTFELFLTIIDGPANY
}			}			NVDLPFMYSITYAAFAIIATLLMLNLLIAMMG
			į	1		DTHWRVAHERDELWRAQIVATTVMLERKLP
}					•	RCLWPRSGICGREYGLGDRWILRVEDRQDLN
						RQRIQRYA
166	1516	A	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS
1						QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF
						NKNMVTERLQNVMVLEQCFSDSSSLYRFLTY
1 1			[[SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI
1 1						WDMRNLATIFLAVVMALLSLHCLAAFKRLE
1						HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER
1						VLYMPSMGYCILFVHGLSKLCTWLNRCGATT
ł						LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS
1						GVQTLPHNAKVHYNYANFLKDQGRNKEAIY
167	1517		2026			HYRTALNNNKAWDYLCWRFRKTLTDLP
1 10/	1517	A	2025	696	71	AAASAASSLTVTLGRLASACSHSILRPSGPGA
						ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY
	J					
1 1						GRLFAVVHFASRQWKVTSEDLILIGNELDLA CGERIRLEKVLLVGADNFTLLGKPLLGKDLV
(RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV
[[TTPQTVLRINSIEIAPCLL
168	1518	Ā	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ
				-		RLQGAARVFMPLQAQVKAKASKPLQMQIKA
			į	ļ		PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS
]	ł	1	1	ļ	i	KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
169	1519	A	2049	1	945	QNLEDREVLNGVQTELLTSPRTKDTLSDMTR
		ļ	-	ļ	-	TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI
	ļ	- 1		j	ļ	EDNSRSKREGLFHENECIVKINNVDLVDKTFA
		i	İ			QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS
ļ .	ļ)	J	J	J	VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
	l		ļ	ì		NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
1			}			SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF
] [ĺ	1	1		TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLO
		ł	l		1	SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
l						ETASLVIARQEGHFLPRELVMFRSQSH
170	1520	A	2050	363	1	PVATHLTKILNSDEHAVVISSAKTLCETVKDF
	ŀ	ł	Ì	ł		VAKVEKTYDKTLENAVVADAVASKCSVLNE
1		Ì		-		KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
<u> </u>			<u> </u>			ESSSEESLGESKEQLGDDVTKPSSQKA
171	1521	A	2055	139	675	IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS
	į	.	ł		1	DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG
						LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL

WO 01/57188 PCT/US01/03800 .

NO. of No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Depide			,				
Sequence				1 .	, , ,		
Sequence		,	[
1914 mg to first manio acid residue of peptide sequence manio acid residue of peptide sequence manio acid residue of peptide sequence manio acid residue of peptide sequence manio acid residue of peptide sequence manio ma	seo-						
maino acid residue of peptide sequence T-Threonine, W-Valine, W-Tyropide sequence T-Threonine, W-Valine, W-Tyropide Sequence T-Threonine, W-Valine, W-Tyropide Sequence T-Threonine, W-Valine, W-Tyropide Sequence T-Threonine, W-Valine, W-Tyropide Sequence T-Threonine, W-Valine, W-V-Tyropide Sequence T-Threonine, W-Valine, W-V-Tyropide Sequence	, -						
Peptide Pept							
Peptide				l			V=Tyrosine X=Unknown *=Ston codon
	ļ		}	1		Sequence	/=nossible nucleatide deletion \=nossible
RESCTITA/PICOSYQDYRNGKCVSCOTSQUE SCPLLG/YADNWADHLRGKDPMTKAFDT	1	1		1			
SCPILGYYADNWKDHLRGKDPPMTKAFFDT		\ 	 	 -	3cquonec_		
AEESFFCMYHYFYDITUNINNVR	· [{				[
172	}	ĺ		1		•	
	172	1522	<u> </u>	2056		261	
	1/2	1522	A	2036	3	301	
H. WRGKEGRRREFMARRILKCLKES				ļ			•
173	1	ł)				
174				L			
AVFLSTVÍLTYTGYJAPWSGRFYSLWDTGYA	173	1523	Α	2060	1	387	
174							
1524	})	}	j l	ļ]	AVFLSVIYLTYTGYIAPWSGRFYSLWDTGYA
174			ŀ		·		KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA
RRIKEBEE.ARI.KYEKEEMERLEIQRIEKEKW					l		G
HRLEAKDLERNBEL EEL YLLERCFPEAEKLK QETKLLSQWKHYIQCDGSPD9VAQEMNT 175 1525 A 2083 139 486 AALTWSQPQEFWFMEMQFIVTDMVTVHWV AESSTVGWLCALFRVTHVGVGATGGVVCG RRVLCGLIPSFAPMPMISLPEGESRKEREVQ RLQFPYLEPGHEPATTILLAFLAAV AESSTVGWLCALFRVTHVGVGATGGVVCG RLQFPYLEPGHEPATTILLAFLAAV RRVLCGLIPLPSFAPMPMISLPEGESRKEREVQ RLQFPYLEPGHEPATTILLAFLAAV RLQFVTHVGVGATGGSEP PQKFLNLLGDITILKGWTGYRGGLDTKNDTT GHISVYTVYQGHEIMFHWSTNLPYSKENKQQ VERKRHIGNDIVTIVFQGEESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEBSYPLFG PPLIPTPYTDHQEFROFILLVKLINGEKATLET PCI GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC CDGAWLAWACWVFGNDFPSPASACSALLG CSVSTACLCVPLCSGSFLAFFRTAAQEGLR RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALQELARCGFILAC RSDLQ RAVSVPLTLAETVASL WAPALGER RAVSVPLTAETVASL WAPALGER RAVSVPLTAETVASL WAPALGER RAVSVPLTAETVASL WAPALGER RAVSVPLTAETVASL WAPALGER RAVSVPLTAETVASL WAPALG	174	1524	Α	2071	74	443	LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE
175			1				RRLKEEEEARLKYEKEEMERLEIQRIEKEKW
1525	ĺ		!	Į i			HRLEAKDLERRNEELEELYLLERCFPEAEKLK
1525	1	ļ	Ì	1			OETKLLSOWKHYIOCDGSPDPSVAOEMNT
AESSTYGWLCALFRYTHYGYGATGHGVYCG RRYLGGILPSPAPMPMSIPEGESRKEREVQ RIQFPYLEPGHELPATILLAFLAAV	175	1525	Α	2083	139	486	
RRVLCGLPLPSPAPMPMSLPEGESRKEREVQ RLQFPYLPPGHELPATTLLAFLAAV 176	1					144	
RI.QFPYLEPGHELPATTLLAFLAAV	1						
1526	1		ł	1			·
FQKFLNLLGDTITLKGWTGYRGGLDTKNDTT GHSVYTVYQGHEIMFHYSTMLPYSKENKQQ VERKRHGNDIVTIVYQGEESSPAKFSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESYPLFG PPLPTPPVFTDHQEFGESSPAFKFSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESYPLFG PPLPTPPVFTDHQEFGELTLVKLINGEKATLET PCI	176	1526	Δ	2002	3	587	
GHSVyTVYQGHEIMFHYSTMLPYSKENKQQ	170	1520	^	2072	١	307	•
VERKRHIGNDIVTIVFOGGEESSPAFKPSMIRS			1				-
HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPVFTDHQEFRDFLLVKLINGEKATLET PCI	1		l	1 1			•
PPL-PTPPVFTDHQEFRDFLLVKLINGEKATLET PC				1			
PCI	1 :		ľ	1 1			
1527	1						9
CDGAWLAWACWYFGNDFPSPASAACSALLG	177	1507		0100	44	405	
CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ 178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHLIFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRFFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVTTQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFF GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPFFPRYLGTSMKAL IQMLEICIGFVVFFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVIMFDDYSATLPLLIVVILENIAVCF	1//	1527	A	2103	44	427	
RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ 178							
RSDLQ	1 .]				
1528							
PSTÄASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRILVKGSKPSQQGRYNMTYERFSSSL LILQVFPNIQNPD 179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFERQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITOGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMOFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			<u> </u>				
PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQIEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	178	1528	A	2104	2	409	
LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PTRSTTPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQIFQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF				i l			
RLQVFPNIQNPD 1529 A 2111 1 312 PTRSSTRPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLACLG					i		
179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSL WLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF				[
MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQITEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVTTQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							RLQVFPNIQNPD
SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQITEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV QFLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVTTQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	179	1529	A	2111	1	312	PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL
PPPLPLACLG	ļ.						MRTESHTGLKKGGNANLVFMLKRNTEPKKG
PPPLPLACLG							SYHFDLERLRAAHILFEREQEHLAPGGISMPL
180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWEVDARRLTRFTGVTTQGRNSLWLS DWYTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWILCFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRNPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	180	1530	A	2116	3	366	
KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVTTQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		-					
QAFASVVCTFHLTACVSLHRIHNSTVV]		l				
181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRNPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	181	1531	A	2117	2	386	·
DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	***	1991	А	***	- .	Jau.	AIDI OOMEMDADDI TETTOMTOODNET VE C
GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1				'		
IS32 A 2123 1 493 RTKTDVYILNLAVADILLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	l			}			`
GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	100	1,520				400	
DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	182	1552	A	2123	1	493	
LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	(1		
RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
	183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ
VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI		Ì			l		RSGNYFVTMFDDYSATLPLLIVVILENIAVCF
					·		VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Velline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVTVLDVND NPPVFTRASYRVTVPEDTPVGAELLHVEASD ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR LAHALDCETQARHQLVVQAADPAGAHFALA PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW PKNNFNGSLVQASYQHEELRREVIMLACSFG NKHCHQQASTLISDWISSNRNRIPLNVRDIVY CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI DVIIHVARNPHGRDLAWKFFRDKWKILNTRI RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYIVHPLWETWA HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP LDEQNRDWQGLLENLHVELTLDEEDSEGPEK EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC AHDKSP
187	1537	A	2158	227	442	FNCFR VASDSFLENSSLLIMILPLRNATQEFIIR PGAVAYTCNPSTLGGWGGWITRSGVRDQPG OHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA CIPONQKMNIWRMKTSKHLQLLSFVLGAVSP AVVVPYMMVLQENGYGVEEGIPTLLMAASS MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV AIVVSSLDW
190	1540	A	2179	64	399	MRLNQNTLLLESFGXXRPYTSEHAPTYHQW MKADELLRWTTSEPLTLEHEYAMQRTWLED AYECTFIVLDAEKRHAQPGATEESCMVGDVN LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR LSYVLFIQERDVHKGMFATNVTENVLNSSRV QEALAEVAAELNPDGSAQQQSKAVNKVKKK AKRILQEMVATVSPAMIRLTGWVLLKLFNSF FWNIQIHKGQLEMVKAATETNLPLLFLPVHR SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS YTHSKGIMHRDVKPLNILCNSPRNKVILADW GLAEFYHPMRKYSVHVATRYYKSPEILLDYE YYDYSLDIWAVGVILLELLTLKLHVFEGGDN EQ
194	1544	A	2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMGS NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE LPVPMGARYIRINPQSWFDNGSICMRMEILGC PLPDPNNY
195	1545	A	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETCD GAVSSLQIVTELQTNYIGKGCDRETYSEKSLQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide insertion KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE MIFKFDGRQGAKIPDGIVPKNLTDQFTTTMW MKHGPSPGVRAEKETILCYSDKTEMNRHHY
196	1546	A	2256	1	396	ALYVHNCRLVFLLRKDFDQADTFRPAEFHW KLDQQALAKVDGQPGKSITRQLQEMPVTIQG ISLKPS FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD
197	1547	A	2259	43	594	NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF KALEESGALLESGTYDGNFYGTPKPPAEPSPF QPDPV QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI
1				,		WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL HCFTESGRGKSWRLCAAILPL
198	1548	A	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATQMFF FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT LMTRKICLQMMMASWMVGFLFSLCIIVTVFN LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM AIFVLSA
199	1549	A	2315	1	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI KRLAFCHSNVLSHSYCVHQDVMKLAYADTL PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	A	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEEDEYDYDYESLSDDNILEDRPENKSCH DQLQFEYKEEM
201	1551	A	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY
202	1552	A	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA. CLAMLLHFLDTYQGLLQEEEGAGHIIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE
203	1553	A	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP
204	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T
205	1555	A	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC EDCSCR

nucleotide sequence USSN (9496) postal minus calculation of the properties of periodic sequence where the properties of periodic sequence provides a minus calculation of the periodic sequence where the properties of periodic sequence properties of periodic sequence properties of periodic sequence properties of periodic sequence properties of the proper	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Sequence 9/14 20/395 2	nucl-	peptide			1 -		F=Phenylalanine, G=Glycine, H=Histidine,
1560	eotide	seq-]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
amino scid residue of peptide residue of peptide residue of peptide sequence T=Threosine, V=Valiae, W=Tryscophan, Y=Trysciae, X=Valosona, Y=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, W=Siop codon, Z=Dispide sequence T=Threosine, V=Valiae, V=Siop codon, V=Dispide sequence T=Threosine, V=Valiae, V=Siop codon, V=Dispide sequence T=Threosine, V=Valiae, V=Siop codon, V=Dispide sequence T=Threosine, V=Valiae, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Valiae, V=Siop codon, V=Dispide sequence T=Threosine, V=Valiae, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Dispide sequence T=Threosine, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon, V=Siop codon,	seq-	uence		09/496	correspondi	to last amino	
Persidue of peptide Sequence Y=Tyrosine, X=Uhdnown, %=Stop podon, Pepsible nucleotide letion, \(\) possible sequence Sequ	uence	İ		914	ng to first	acid residue	
Persidue of peptide Sequence Y=Tyrosine, X=Uhdnown, %=Stop podon, Pepsible nucleotide letion, \(\) possible sequence Sequ	ì			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Peptide	1		1		residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
					peptide	·	
1556	ĺ		[sequence		nucleotide insertion
THIPCPVSSTINDTPDQIWSVOSZIMGTIGG MGANASTSPRCVBLSGORKKWIQVPLASIV	206	1556	A	2406		485	DLSPDSREDHPOGHRRLLPKRPVRGSLMPGH
MGANASTSPRCWDLSSGNKWIQVPILAST 207			1				THIPCPVSSTTNDTPDOIWVSVGSLRMGTGG
	1		1			İ	
1557 A 2409 289		ļ	İ				
HTIM/LIKENTILS	207	1557	A	2409	289	418	
208	1		1			} ···	
MLNI.GLGSMIGTMAGITTPIIDTEKVPKEMFT GGCVYPAL VGLIFVQRSGNYTMFDDYSA TLPLTLIVILENIAVAWIYGTKKFMQELTEML GREYTRYTYTMWKFVS 209	208	1558	A	2413	64	492	
GGCCYPAFLVGLLFVQRSGNYFVTMFDDYSA		1000			,	'/-	· •
T.P.I.T.I.VILENIAVAWITGTKKFMQELTEML GFEPYRFYFTMYKFVSP	1						1
	!						I
209 1559 A 2417 3 877	1 .		ŀ	ļ		Į	
NASNLDKVI.TDIKADKDQANDGI.SSALILIY LDSARNLPIRYKTNEPVWEEPITFIHNPKRQ DLEVEVRDEQHQCPLGNIKVPLSQLLITSEDM TVSQRPQLGNSGPNSTIKMKIALRVHLEKRE RPPPHQHSAQVKRPVSKEGKTSIKSHMSG SOPGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLIBUGGSSSSLLASPGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTILGQSPLGQI QLTIP 210 1560 A 2422 35 456 REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNOSLAATIDHRPCSAPLEPRIQASRNQRW GAVRAARSI.TDIAPPASPQVHETPIDASQTOK VEPASKSRTTPELQAKVSHSRERALSTMDATP HHAQPQRGEG 211 1561 A 2431 1 764 RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLMAAMFE ANGGGVYLAPPVLWSKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHPTKTRHNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFNSRM EENSSGNTMSASSMPTIERKKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YGNGTKKEDDSLTFFAVAKRDHYSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLBLTVLAERIVRLTWIPGDANNRPI TDYDCQLEEHQ 213 1563 A 2445 1 1294 MSSIGCLWSYRSSQDGLTAEKSGPEKPHGT WLMPELIFIKEQULELLVLEQFTSLFELQTWV QQNNPESGESSVTLEDLERDEPDPQQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEKSQG LPQEPSFRGISSHESNLVWKQGSATGEKLRSP SQGGSSQVIFTNKSLGKRDLYDEARCLLT TDSIMCGKVPPEEPTYRCDVCGHISFKQHSSLT QHQRHTGEKPYKCNQCGKAPSLRSYLIHQR HSGEKAYECSECGKAPNQSSALIRHRKHTG EKACKCNECGKAPSQSSYLIHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTFTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY ECNECGKTTSQSSKLIRHQRHTGERPY	200	1550		2417	3	077	
LDSARNLPIRYKTNEPVWEENFTFFINPKRQ DLPEVENBEDHOQPCIONTSYLSQLITSEDM TVSQRPQLGNSGPNSTIKMKIALRVLHLEKRE RPPDHQHSAQVKRPPVSKEGRKTSKSHMSG SPGPGSSNTAPSTIVIGGSDKPGMEEKAQPPE AGPQGLIDIJGRSSSSLLASPGHISVKEFTPSIA SDISLPIATOPELRQRLRQLENGTTLGQSPLGQI QLTIP 210 1560 A 2422 35 456 REFAASDLEPFTPTDPISPEATTQPSCICRQRA AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRTFTELQAKVSHSRERALSTNDATP HHAQPQRGEG RYSQKLIQHTACQLLRTYPAATRIDSSNPPP LMFWLHGIQLVALNYQTDDLPLHAMAMFE ANGGCGYVLKPPVLWDKNCPSNSMGSPCIE VDVLGMPLDSCHFTKPHRNTLNFMWNEQF IFVHFEDLVFLRFAVVENNSAVTAQRIPL KALKRGYRHLQLRNLHNEVLEISSLFNSRM EENSSGNTNSASSMPTEEKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPPNDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLITTAVAKRCHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWPGDANNRPI TDYDCQLEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT VLMPELLIPKEQULELLVLEOFLSILPELQVW QQHNPESGEESVILLEDLEREDDPQQVPAS PQGPAVPWKDLTCLASAGSETDHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEKSQG LPQFSFRGISHESNLWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEARCLLIT TDSIMCQKVPPEERPYRCDCGSFKGHGSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIHQR HISGEKAYECSECGKAFNQSSALIRHQRHTGERPY ECNECCKITSQSSKLIRHQRHTGERPY ECNECCKITSQSSKLIRHQRHTGERPY ECNECCKITSQSSKLIRHQRHTGERPY ECNECCKITSQSSKLIRHQRHTGERPY ECNECCKITSQSSKLIRHQRHTGERPY ECNECCKAFT ENTRYSVQC	209	1339	^	2417	3	0//	
DLEVEVRDEOHQCPLGNILKYPLSQLITSEDM TVSQRFQCLGNSGPNSTKMLALRVLHLEKRE RPDHQHSAQVKRPSVSKEGRKTSIKSHMSG SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSSLLASFGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTP	1		1			j	
TVSQRFQLGNISGPNSTIKMIALRVLHLEKRE RPPPHQHSAQVXRPSVEGRKTSIKSHMSG SPGPGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSSLLASPGHISVEFTPSIA SDISLPIATQELRQRLAQLENGTTLGQSPLGQI QLTIP 210 1560 A 2422 35 456 REFAASDLEPFTPTDQPISPEATTQPSCIKRQRA AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTVG VEPASKSRTPPELQAK YSHSRERALSTMDATP HHAQPQRGGB 211 1561 A 2431 I 764 RRYSQKLIQHTACQLERTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLITVSQQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPHRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIPL KALKRGYRHLQLRNLHNSVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 I 411 GIRGTGHLGCPNDDPSLTLTVSWVMEDKPI YGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATFTNRLGALPKG RANRTRDLELTYLAERVRLTWIPGDANNRPI TDYDCQEHQ 213 1563 A 2445 I 1294 MSSIGCLWVSRSSQIDGLTAEKSGFEKPHGT WLMPELHPKEQULELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQVPAS PQGPAVPWKDLTCLRASQESTDHLQPLKTQ LKSWKPCLSPKSLCEMSTATKEGISEKSGG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGSFSQVIFTNKSLGKRDLYDEARECLBLT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYYKCNQCGKAFSLSSVLIHQRHTGERPY ECNE LKSWKPCLSFKSSCCBAFSQSSYLIHQRHTGERPY ECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHGRHSGEKPYECSECGKA FSLSSNLIRHQRHISGEKPYECSECGKA]						
RPPDHQHISAQVKRPSVSKEGRKTSIKSHMSG SPGGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSLLASFGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTP 210 1560 A 2422 35 456 REFAASDLEPTPTDQPISPEATTQPSCIKRQRA AGNRGSLAATIDHKPCSAPLEPRIQASSRNQRW GAVRAAGSLIDIAEPASPQVHETPDDASQTQK VEPASKSRTPBELQAKVSHSRERALSTMDATP HHAQPQRGGG 211 1561 A 2431 I 764 RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPHLNAAMFE ANGGGGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSISMGSPCIE VDVLGMPLDSCHFRTKPHRNTLNPMWNEQF LFFVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYSHLQLRNLHNEVLEISSLFINSRRM EENSGGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 I 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YGNGTKKEDDSLTIFAVAKRDHYSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANNTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 I 1294 MSSIGGLWVSRSSQIDGLTAEKSGPEKPHGT WLMFELHPKEQLELLULEQFISLPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSFKSDCENSSTATKEGISSEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSTSOVIPTNSLGRLYDLAGRCLLIT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYXCNQCGKAFSLRSYLIHIQR HISGEKAYSCSCCGKAFPQSSALIRHQRHITGERPY ECNECKETSQSSKLIRHQRHTGERPYECNE CGKAFRQSSELITHQRHSGEKPYECNE CGKAFRQSSELITHQRHSGEKPYECNE CGKAFRQSSELITHQRHSGEKPYECSECGKA FSLSSNLIRH	1		ļ				
SPOPGGSNTAPSTPVIGGSDKPGMEKAQPPE AGPQCLHDLGRSSSSLLASPGHISVKEPTPSIA SDISLPIATGELRQRLAQLENGTITLGQSPLGQI QLTIP 210 1560 A 2422 35 456 REFAASDLEPFTPTDQPISFEATTQPSCIKRQRA AGNPGSLAATDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDLAEPASPQVHETPIDASQTQK VEPASKSRTTPELQAKYSHSRERALSTMDATP HHAQPQRGEG 211 1561 A 2431 I 764 RRYSGKLJOHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSQNVCPSNSMGSPCIE VDVLGMPLDSCHFRIKPHRNTLNPMWNEQF LFFHYHEDLVFLRFAVENNSAVTAQRIPL KALKRGYRHLQLRNLHNEVLEISSLFNSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG CVPG 1562 A 2436 I 411 GIRGTTGHLGCPNDDPSLTLIVSWVMEDKPI YGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEHQ MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQUELLVLEQFLSILPELQIWV QQHNPESGEESVTLLEDLEREFDDPGQVPAS PQGPAVPWKDLTCLRASQESTDHLQPLKTQ LKSWKPCLSFKSSCENSTATKEGISESKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSPSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKYYKCNQCGKAFSLSSVLIHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHGRHISGEKPYECSECGKA FSLSSNLIRHQRH	ì						, ,
AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTIP 210 1560 A 2422 35 456 REFAASDLEPFTPTDQPISPEATTQPSCIKRQRA AGNPGSLAATIDHAPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSSERALSTMDATP HHAQPQRGEG 211 1561 A 2431 1 764 RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAYVSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTPHIRNTLNPMNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIPL KALKRGYRHLQLRNLINEVLEISSLFINSRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTINGLGALPKG RANRTRDLELTVLAGPLSTURTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQLELLVLEQFLSLFEELQIWV QQHNESGGESVTLLEPLERFDDPGQQVPAS PQGRAVWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKGISEEKSQG LPQEPSFRGISSHESNLWWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEARCLLT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLSSYLIHQR HISGEKAYECSGCKAFNQSSALIRHKKHITG EKACKCNECGKAFNGSANITN.SLDRSG S	į į						,
SDISLPIATQELRQRIRQLENGTTLGQSPLGQI QLTIP 210 1560 A 2422 35 456 REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNPGSLAATIDHRPCSAPLEPKIQASRNORW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSRERALSTMDATT HHAQPQRGEG 211 1561 A 2431 1 764 RRYSQKALQHTAQQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTRPHRNTLNFMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHQLKNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG CYPG CYPG CYPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVGQAPTPTNLGGLPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ MSSIGCLWVSRSSQDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGESVTLLEDLEREFDDPGQVPAS PQGFAVWKDLTCLRASQESTDHHQPLKTQ LKSWKPCLSPKSDCENSETATKEGISSEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGFSQVIFTNKSLGKRDLYDEAERCLLLT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQHITGERPYKCNOCGKAPSLRSPSLIHIQR HTGGEKAYECSECGKAPNQSSALIRHRKIHTG EKACKCNECGKAFSQSSTVLHIQRIHTGERPYECNE CGKAFRQSSELITHQRIHTGERPYECNE CGKAPQSSELITHQRIHTGERPYECNE CGKAPQSSELITHQRIHTGERPYECNE CGKAPQSSELITHQRIHTGERPYECNE CGKAPRQSSELITHQRIHTGERPYECNE CGKA				l			
OLTIP	[
210							
AGNIPGSLAATIDHKPČSAPLEPKŪQASRNQRW GAVRAAESLITDIAEPASPQVHETPIDASQTQK VEPASKSRITPELQAKVSHSRERALSTMDATT HHAQPQRGEG 211 1561 A 2431 1 764 RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWHGQUVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLITVSQNVCPSNSMGSPCIE VDVLGMPLDSCHFRIKPHRNTLNPMWNEQF LFHVHEEDLVFLRFAVVENNSSAVTAQRIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTINRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ VQHNPESGEESVTLLEDLEREFDDPQQQVPAS PQGPAVPWKDLTCLRASQESTDHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGFSQVIFTNSLGKRDLYDEAERCLLT TDSIMCQKVPPERPYRCDVCGHSFKQHSSLT QHQRHTGEKPYYKCNQCGKAFSLRSYLIHQR HSGEKAYSCSECGKAFNQSSALIRHRKHTG EKACKCNECGKAFSQSSYLIHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECSECGKA FSLSSNLIRHQRHTGERPYECSECGKA SSLSSNLIRHGRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHGRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHGRHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRGSSELITHGRHTGERPYECNE CGKAFRGSSELITHGRHTGER	L						
GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSRERALSTMDATP HHAQPQRGEG 211 1561 A 2431 1 764 RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPJHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAYYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPHRNTLNPMWNEOF LFHVHEDLVFLRFAVVENNSSAVTAQRIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GRGTTGHLGCPINDDPSLTLTVSWVMEDKPI VIGNGTKKEDDSLTIFAVAKRDHYSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQLELLULEQFLSILPELQIWV QOHNPESGESVTLLEDLEREDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSFKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHSNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLLLT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGERPYRCKNQCGKAPSLRSYLIHQR HSGEKAYECSECGKAFNQSSALIRHRKHTGE EKACKCNECGKAFSQSSKLIRHQRIHTGERPY ECNECKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHTGERPYECNE CGKAFRQS	210	1560	A	2422	35	456	
211 1561 A 2431 1 764 RRYSQKLIGHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPIHRNTLNPMWNEOF LFHVHFEDLVFLRFAVVENNSSAVTAQRIPL KALKRGYRHIQLRNLHNEVLEISSLFINSRIM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTINRLGALPKG RANNTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILEILVLEQFLSILPELQIWV QQHNPESGEESVTLLEDLEREFDDPQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSTSQVIFTNKSLGKRDLYDEAERCLLLT TDSIMCQKVPERPPYRCDVCGHSFKOHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIHQR HSGEKAYECSECGKAFNQSSALIRHRKHITG EKACKCNECGKAFSQSSYLIHQRIHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHGRIHSGEKPYECSECGKA FSLSSNLIRHQRHSG SPMVPSYETSVSPQANTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR				٠ ا			
HHAQPORGEG	1						
1561							
LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQDIVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPHRNTLNPMWNEQF LFFYHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGQATPTINRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGESSVTLLEDLEREFDDPQQVPAS PQQPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQQGSFSQVIFTINKSLGKRDLYDEAGRCLLT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHORIHTGEKPYKCNQCGKAFSLRSYLIHQR HISGEKAYECSECGKAFNQSSYLIHGRHTGERPYECNE CGKAFRQSSELITHQRHSGEKPYECSECGKA FSLSSNLIRHQRHSG 214 1564 A 2461 1 615 GIPGSTISSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITINLSLDRSG SPMYPSVETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR							
ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPHENTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELIPKEQILELLVLEQFLSILPELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIHQR HISGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHTGERPYECNE CGKAFRQSSELITHQRIHTGERPYECNE CGKAFRQSSELITHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG STRLVSVQEDAGKSFARNRSASITNLSLDRSG SPMVPSYSTSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	211	1561	Α	2431	1	764	
DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFTIKPHIRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVG GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI Y1GNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTINLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGSFSQVIFTNKSLGKDLYDEAERCLILT TDSIMCQKVPPERPYRCDVCGKAFSLRSYLIHQR HSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAOKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDBERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR							LMFWLHGIQLVALNYQTDDLPLHLNAAMFE
VDVLGMPLDSCHFRTKPİHRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNILHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTINRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDHHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSMCQKVPPEERPYRCDVCGHSFKQHSSLT QHRIHTGEKPYKCNQCGKAFSLRSYLIHQR HISGEKAYECSECGKAFNQSSALIRHKIHTG EKACKCNECGKAFSQSSYLIHQRHTIGERPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHSGEKPYECSECGKA FSLSSNLIRHQRHSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	1	· .					ANGGCGYVLKPPVLWDKNCPMYQKFSPLER
LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTIRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQLLELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIHQR HSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRHTGERPYECNE CGKAFRQSSELITHQRIHTGERPYECNE CGKAFRQSSELITHQRHTGERPYECNE CGKAFRQSSELITHQRHSGEKPYECSECGKA FSLSSNLIRHQRHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR							DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE
KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG	ì	1					VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF
EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG 212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTFL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	.l	١ .					LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL
212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIHQR IHSGEKAYECSECGKAPNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRHTGEKPYE CNE CGKAFRQSSSYLIHQRHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECNE CGKAFRQSSELITHQRIHSGEKPYECNE STSLSSNLIRHQRIHSG SPMYPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR							KALKRGYRHLQLRNLHNEVLEISSLFINSRRM
212 1562 A 2436 1 411 GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHINLDKGYLTVLGEQATPTINRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKAFSQSSYLIRHQRHTGERPYECNE STSLSSNLIRHQRIHSG SPMYPSYETSVSPQANRTSYNLSLDRSG SPMYPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR]						EENSSGNTMSASSMFNTEERKCLQTHRVTVH
YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLLLT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR HISGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	1 1						
YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRHISG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	212	1562	Α	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI
TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVITLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR HISGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGERPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR]						
RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR IHSGEKA YECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	ļ l					. !	
TDYDCQIEEHQ 213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR							
213 1563 A 2445 1 1294 MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	j j			J			
WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	213	1563	A	2445	1	1294	
QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR							
PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR				· i	ļ]	
LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR					Ì		
LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR		ļ					
SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	[1		. 1	ſ		
TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR IHSGEKA YECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR					l		
QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	[. 1		
IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR				·	i	' I	
EKACKCNECGKAFSQSSYLIIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR							
ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	[[Ĭ	ĺ	ĺ		ĺ	
CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG 214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	1 1	Į	Į				
214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	[1	1	ļ		ļ	
214 1564 A 2461 1 615 GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR		1	- 1	l		Ì	
STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	1214	150					
SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	214	1564	Α	2461	1	615	
LDSVQLKDLWKKICHHSSGMEFQDHRYWLR	ļ ļ		J	j	J		
	1 1	ı					
THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ	1	ŀ	ļ	J		ļ	
							THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ

SEQ II NO: of		Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide] ===	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	[USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ļ	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]	J		J	peptide		/=possible nucleotide deletion, \=possible
		Ì	ļ	sequence		nucleotide insertion
 		 	├──	- Sequence	 -	AMVDGRWLDCVSHHDQLFRDEYALYRPLQV
1	1	ļ	ļ	1	ļ	LFSVYCOLECSKLIL
215	1565	A	2464	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGO
213	1303	^	2404	-	2732	HGPGRHGRRVCSSQDSMADVFVHLRTAWPT
]	1		1	j		CSLISGOHGPGESVSYEDDDIPAPASLLHVNA
						AAPALTNPTAPVLCTAPNNTAOKEKVPSGMR
						ORPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
j]		1]		TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
	1					STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1				•	
ŀ	1	1				TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
1						ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH
	i		1			
ł			ļ			ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	į:					TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
	ŀ		1			ATVPGVRISSCTPDLTCAVSIHATVPGVRISSC
	}		l .			TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
						ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
		1				TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH
l	1					STVPGVRISSRTPDLTCAVSIHATVPGVHISSC
İ	İ				·	TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH
1						STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT
1	1		ļ			PDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS
İ						TVPGVHISSCTTDLTCAVSIHATVPGVHISSCT
1	-	l				PDLTCAVSTHTTVPGVRISSRTPDLTCAVSIHS
1	1	1				TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT
						PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA
		<u> </u>				TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT
l	- 1	1				PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS
ĺ]			TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT
ĺ	.l		1			PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH
l	Ì		1			STVPGVHISSRTPDLTCAVSHATVPSVHISSC
						TPDLTCAVSIHSTVPGLLTSVSQTSTG
216	1566	A	2477	1	414	FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK
	ł	}				AVEVATVVIQPTVLRAAVPKNVSVAEGKELD
						LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS
İ			1			RVLARLDRDFLVHSSPHVALSHVDARSYHLL
	}			_		VRDVSKENSGYYY
217	1567	A	2480	2	460	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY
1			ŀ			MQVLVCQHECVRELATRPGRLSPIENFLPLHY
1	1					DYLQFAYYRVGEYVKALECAKAYLLCHPDD
l						EDVLDNVDYYESLLDDSIDPASIEAREDLTMF
						VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
ł	1					SANLLQLVRSSGDIQEGDLVEVVLSASATFED
ĺ				İ		LQIRPHALTVHSYRAP
219	1569	A	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLO
				-		CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
						DCKYECMWVTVGLYLQEGHKVPQFHGKWP
						FSRFLFFQEPASAVASFLNGLASLVMLCRYRT
]	Ì			ļ	FVPASSPMYHTCVAFAWVS
220	1570	A	2498	1	1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA
	15/0		0,470	- }		APKIPRLVQATPAFMAVTLVFSLVTLFVVDH
	1			. 1	. 1	
			, 1			HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSOLOVLGDHLGNTNADIOMV
				,		
					ļ	, , , , , , , , , , , , , , , , , , , ,
						KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL
						KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL
						KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS
						KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	noc	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	!	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ	i .	peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence	<u> </u>	nucleotide insertion
						FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM
						VTAQTQKANGRLDQTDTQIQVFKSEMENVN TLNAQIQVLNGHMKNASREIQTLKQGMKNA
						SALTSOTOMLDSNLOKASAEIORLRGDLENT
			·			KALTMEIQOEOSRLKTLHVVITSOEOLORTO
221	1571	A	2501	3	500	RVRLNNDGLSPLMMAAKTGKIGIFOHIIRREV
						TDEDTRHLSRKFKDWAYGPVYSSLYDLSSLD
1	l					TCGEEASVLEILVYNSKIENRHEMLAVEPINE
						LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT
1		1	İ			AYYQPLEGTPPYPYRTTVDYLRLAGEVITLFT
1222	1672		2500		205	GVLFFFTN DANOGRAL AMOREM PRICE TELL TELL TELL TELL TELL TELL TELL TE
222	1572	A	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
						ARHVRDKEEEVDLVMQKVESLRQELRRTER AKKELEVHTEALAAEASKDRKLREQSEHYSK
		1		'		QLENELEGLKQKQISYSPGVCSIEHQQEITKL
						KTDLEKKS
223	1573	A	2544	2	412	NDPALISNESAAVVHTIVNETLESMTSLEVTK
						MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS
						EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID
						KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ
	1.50			401		LTHCSLSAA
224	1574	A	2552	401	1	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM
		1				CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI
						QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA
						VK
225	1575	A	2563	724	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE
						GLGEEEKEAGKKKKKQEEKEKEKGAVYSR
j						VARICKNDMGGSQRVLEKHWTSFLKARLNC
1						SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT
						QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP
						DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK
1 1						TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT
226	1576	Ā	2571	449	3	KTRVRYRLTAISVDHSAGPYH EGVLFVYGNYVGDVMNFEMAAEMAOEVAIP
220	1370	^	23/1	7777		TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV
			i			AGAACDRGMSLEACEAVTRKANRRTYTMG
						VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE
]]						RGVIREKMMPADAIVDHIMDRIFS
227	1577	A	2575	3	1197	VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS
						EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
						ELWMKAMLDAALVQTEPVKRVDKITSENAP
				ļ		TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE
				i		KKALEAEKYGFQKDGQDRPLTKINSVKLNSL
]					PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIORTNSMOO
						LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS
	ļ				(HRAQIMARYPEGYRTLPRNSKTRPESICSVTP
						STHDKTLGPGAEEKRRSMRDDTMWQLYEW
] }						QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT
						MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI
						QRGDVTIDRRHRAHHPKVK
228	1578	A	2583	3	330	LPFLGLGSVLPQGMVMASPEMNPTICSVFEA
					·	HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE
		- 1	ł		•	KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD
220	1570		2500		440	PTMGIKPHLWWVAA
229	1579	A	2589	1	448	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR FCVAPNICKCKPGYIGSNCOTAL CDPDCKNH
[]		. [ĺ	Ì		ECVAPNICKCKPGYIGSNCQTALCDPDCKNH GKCIKPNICQCLPGHGGATCDEEHCNPPCOH
L						GROWN MICCOLL GROWN CONEENCHT COM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide			USSN	location		
	seq-	İ	09/496		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence			correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	į		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	Ì	l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1 :	ĺ	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ '	1	l		peptide		/=possible nucleotide deletion, \-possible
		<u> </u>		sequence		nucleotide insertion
	·					GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
						ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	A	2593	2	138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI
	ŀ		,		1	LYLLYLLRNAFFLR
231	1581	A	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER
ľ l					i	WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	A	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
1					***	GRARRRTPTCEPATPLCCRRDHYVNFQELGW
l i						RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS
1				1		FHSAVFSLLKANNPWPGRTSWCVPTARRPLS
						LLYL
233	1583	A	2601	184	403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
23	1505	A	2001	104	703	YNGKETTLGDMTGKCKSWITPCPEEKVNVLO
						NSIPYWERIT
234	1584		2614	178	225	
234	1384	A	2014	1/8	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
025	1506		2616		004	DGSKSSDDQKIISYLWEKTQ
235	1585	A	2616	2	896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
						AKFLNVEAAMVFGMGFATNSMNIPALVGKG
)						CLILRDEVNHTSLVLGARLLGATIGIFKHNYA
Į į						QSLEKLLRDAVIYGQPRTRRAWKKILILVEGV
						YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI
						GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
						FGASGGYIAGRKARILSPPACLVPNTGSHSLH
						RLTRDLQMNEAMVALVTDRLQGWNSGEGN
İ						WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
1						AEQIIRSLKLIMGLDGTTQ
236	1586	Α	2621	ì	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS
						ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
1						WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
						KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
]						A
237	1587	A	2628	398	1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
1					-	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
l i						GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV
1 1						QDPEPPNV
238	1588	A	2631	1	1104	WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
20	1366	^	2031	•	1104	ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
						WQPCSRTCGGGVQKREVLCKQRMADGSFLE
!						
1 /	ŀ					LPETFCSASKPACQQACKKDDCPSEWLLSDW
						TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
] }						TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI
i I					İ	AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA
						VVLRCPARRVRKPLITWEKDGQHLISSTHVT
1 1		ı				VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
						VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
1 1		ı	- 1		l	KEALQTHKHQNGIFSNGSKAEKRGLAANPGS
<u> </u>						RYDDLVSRLLEQGAPCSSSKKKN
239	1589	A	2636	1	678	MKPDNILLDEHGHVHITDFNIAAMLPRETQIT
]						TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
		.[[ſ	SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
				ľ	ļ	TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ
[[{	İ	Í	i	ĺ	LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK
			Ţ	l		GRLNCDPTFELEEMILESKPLHKKKKRLAKK
(/	ľ			i	ľ	EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE
	- 1	l		ł	l	KVNRDCI
240	1590	A	2639	389	3	ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
•	/			707	-	LWQNFAREIEEHVFTLYSKNIKKYKTCIRSKV
	}	J	ļ		j	ANLKNPRNSHLQQNLLSGTTSPREFAEMTVM
<u> </u>						WATER I GOTTLE AND I 19 LYELYEMI AM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EMANKELKQLRASYTESCIQEHYLPQVIDGTL
					 	Y
241	1591	A	2640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	A	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT STTAMVMPIVEAVLQELVSAEDEQLVAGNSN TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK SPLMISQACI
243	1593	A	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT QNSETV
244	1594	A	2650	1		MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLLMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNPSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE VVSDPASV
245	1595	Α	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
246	1596	A,	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
247	1597	A	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	A	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
249	1599	Ā	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK AIDLSRNQFQDFPEQLTALPALETINLEENEIV DVPVEKLAAMPALRSINL
250	1600	A	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
251	1601	A	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
	;			amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
						QKRKKKAPDHSSGRKEELVTTHTVDKLETKK PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH
						RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI GAGDCL
252	1602	Α	2697	421	1	PQKSHSGAYQCFATRKAQTAQDFAIIALEDG TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT
						VTWALDDEPIVRDGSHRTNQYTMSDGTTISH MNVTGPQIRDGGVYRCTARNLVGSAEYQARI NVRGPPSIRAMRNIT
253	1603	A	2698	65	401	ACCOWRRTLIPAKSTTVSCTISTPHHPFRGSYS FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL
						PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG AAFTNNIASSTIIL
254	1604	A	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ RVRGPWEAGPGVGY
255	1605	A	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
						AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA DKETLENMMQRHEEEAHEKGKILSEQKAMIN
						AMDSKIRSLEQRIVELSEANKLAANSSLFTQR NMKAQEEMISELRQQKFYLETQAGKLEAQN
						RKLEEQLEKISHQDHSDKNRLLELETRLREVS LEHEEQKLELKRQLTELQLSLQERESQLTALQ
	•					AARAALESQLRQAKTELEETTAEAEEEIQALT VGLGSNIFRLLKASARMSVELALSILAHP
256	1606	A	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE LLKESGNQVLKNGNFSLAIRKYDEAIOILLOL
	!					YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ
257	1607	Α	2702	2	399	PYEAARMFFEGLR FVESASSRPPGCFSGDGRFWLVSEGSRRGWD
		•-		_	333	FNPSFSFLDPRYSVGGDENIGTVTTLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE
						DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL
258	1608	A	2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI
						MTRKELLTVYSSEDGSEEFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
						QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH
						SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPGQGFGESMYLLSD
						KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR
i			•			LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD
259	1609	A	2721	1	403	ARAFFAQDGVQSLPTQKWWGDMARR VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR
					.05	GGFSQREMVTGERSPSPEEEEEEEEGFGERA
						SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE
260	1610	A	2728	1	477	GGLRVRLP LLGGDLRYHLQQNVHFTEGTVKLYICELALA
						LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM
						DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL
261	1611	A	2730	3	547	LTITOFILVLYRYYRSPLVQIYEIEQHKIETWR
L						EIYLQGCFKPLVSISPNDSLFEAVYTLIKNRIH

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
	i			amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		RLFVLDPVSGNVLHILTHKRLLKFLHIFGSLLP RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL DIFVDRRVSALAVVNECGTHPQDERLGLGW GLGEPGSEERLFPAAITSR
262	1612	A	2733	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA GRLVKLSLANNNLVGVHEDAFETLESLQVLE LNDNNLRSLSVAALAALPALRSLRLDGNPWL CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM ESRRISLRACRRPASRV
263	1613	Α	2736	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWQF LWGLNGNFNFFKEPWGGRNNHAKGFRTTW ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA RLISPLVNLPQSPGGLEFQYQAT
264	1614	A	2738	2	245	RAMLKCLREGQPPPSYNWTRLDGPLPSGVRV DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH DTVDVLDPPEDSGKQVDL
265	1615	A	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV LAAVETTVLVLIFAVSLLGNVCALVLVARRR RRGATACLVLNLFCADLLFISAIPLVLAVRWT EAWLLGPVACHLLFYVMTLSGSVTILTLAAV SLER
266	1616	A	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP V
267	1617	A	2760	434	714	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
268	1618	A	2762	1	405	IACTFCGQDEWSPERSTRCFRRRSRFLAWGEP AVLLLLLLSLALGLVLAALGLFVHHRDSPL VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE LPLSWAE
269	1619	A	2772	3	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN LSKNLSFSEFCFDVSY
270	1620	A	2789	1	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA VEQIAKAEETHSSLSQELQARLQTVTREKEEL LQLSIERGKVLQNKQAEICQLEEKLEIANEDR KHALERFEQEAVAVDSNLRVRELQRKVDGIQ KAYDELRLQSEAFKKHSLDLLSKERELNGKL RHLSP
271	1621	A	2795		568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
272	1622	Ā	2797	8	523	VKGIQEKQVFSNTKDSKQEITQNKSFFSSVKE SQRDDGKGALNIVEFLRKREELHQILSTVKQP KCMQGKYAGAMESEPCVCTEADFDCDYGYE
						RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST GYRKVVSNNCTDGVREQYTAKPQKCPGKAP RGLRIVTADGKLTAEQGHNVTLMVQLEEGD VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
273	1623	A	2801	72	395	YQNXGIXRXTVQVDNSLGS HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS DEEGQNQLGIRIHQDIPLPPRRRELPALRTING KADSLNVSRNSVMQELSELEKQIQVIRQELQL
274	1624	A	2805	168	320	AVSRKTELEEYH ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE IFIARNGVVGETLTHCKRV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghttamic Acid,
nucl-	peptide	~~~	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Ghrtamine, R=Arginine, S=Serine,
	ļ	j	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1		peptide	Sequence	/=possible nucleotide deletion, \=possible
1				sequence	ļ	nucleotide insertion
275	1625	A	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
2/3	1023	1 ^	2012	208] 321	MGKIIFQ
276	1626	A	2813	41	266	
270	1020	Α.	2013	41	200	AGRSLHGAGDRAWVGISPTDWSPKVVELCK
		1				KYQQQTVVAIDLAGDETIPGSSLLPGHVQAY
277	1627	_	2817		410	QVGPVRRNGEAGPG
211	102/	Α	2817	3	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
1.		l				LFISYLHTPKHKQHEVLQAMGSILGITGEEME
l i		1	1			PLFQEEHGTATRWMTGWLEGGSKSVPKTPL
		l				GLNQQPALNGSFSELFVKFLKTESLSSTLPTX
	1/20		0001	222	L	LPPHNSPGKIK
278	1628	A	2821	238	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE
1						VKLRLLLHLEELQMEHDIRHYDLESVPMTWD
<u> </u>		<u> </u>	L		ļ	PVDQNPRLV
279	1629	A	2822	342	1	PLIPANLPAHSNPLQPLPSLPHPFLPATHKFPT
		ĺ				TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLQS
j			i			CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS
<u></u>						VPSQRHPTXPPPAS
280	1630	Α	2825	307	77	PSMVWSYHWGVKQKRLALCVFSFEEGGRRK
1 1						CGQYWPLEKDSRIRFGFLTVTNLTGAVGEPG
l l						VAFQCDGQRRREPTC
281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASQEGGDVLGAR
1 1						QDCTPSLKSLVATGNLLDLEETAKAPLSTVSA
}						NTTNMDEVPRPQALSGSSVVWVSGCVASRS
						VILSLTSG
282	1632	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWOVFV
						TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
						YNYPVLLPLLDDLFKVHKLKPNLKWROAFDS
						YLKTLPPYYL
283	1633	A	2835	462	148	VSPALSLTPTIFSYSPSPGLSPFTSSSCFSFNPEE
}						MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
						PLOCOMHPEESTOFSIKLOPPPVGRKNRERVE
			· •			SSEESAP
284	1634	A	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ
		• •	2000	-	,	DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL
f I						KSLAETVLNFPLDKSLLLRCSNWDAETLTED
	ì					QVIYAARDAQISVALFLHLLGYPFSRNSPGEK
					!	KR
285	1635	A	2843	20	271	PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL
	1000	**	2043		~/1	VCDRVSEDGINRQQAQEWCIKHGFELVELSP
1 1	ł					EELPEEDGKCLCVRRKYGTYI
286	1636	Ā	2845	197	278	TAEDVLTVAYEHGVNLFDTAEVYAAGK
287	1637	A	2851	2	427	FVAEVRREWAKYMEVHEKASFTNSELHRAM
201	103/	^	2031	-	441	
	1			1		NLHVGNLRLLSGPLDQVRAALPTPALSPKDK
[[ĺ		1	ĺ	(AVLQNLKRILAKVQEMRDQRVSLEQQLRELI
ļ [ļ		l		QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL
200	1636	_ ; _	2050		420	KVYLEQNLAAQDRVLCALT
288	1638	A	2859	2	469	FVNLGILTCIECSGIHREMGAHISRIQSLELDK
	· [į	l	LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
		1			1	TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA
	1			ł		KLNELLEAIKSRDLLALIQVYAEGVELMEPLL
			1			EPGQELAETALHLAVRTADQTSLHLVE
289	1639	A	2861	2	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
	. 1	ļ	ļ		j	DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
·	ĺ		Ì		ļ	DSGLWRMHLMEGELPASMSGSCGACINGKL
	ì	ŀ	ł		ł	YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
	, [ļ		ITDFEGQPPTPRDKLSCWVYKDRLIYFG
290	1640	A	2868	1	378	FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI
	ļ	1	.	.	`	SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF
	1	ļ	j	l		PDCASCLQAQDPLCGWCVLQGRCTRKGQCG
				ــــــــــــــــــــــــــــــــــــــ		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						RAGQLNQWLWSYEEDSHCLHIQSLLPGHHPR QE
291	1641	A	2870	1	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL PFTPKSIRSHFQHVFVIVKVHNPCTENVCYSV GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL AKVINAENAAHKSEKFRAMATRTRQEYLKD LA
292	1642	A	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI PPPPAVPYSPRYVAVHCHGMLVSCWCHL
293	1643	A	2878	1	427	REKEEEVEEEEDKVVKETEKEAEQEKEEDSL GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF LSPEKLTAENRYYCESCASLQDAEKVVELSQ GPCYLILTLLRFSFDLRTMRRRKILDDVSIPLL LRLPLAGGRGQAYDL
294	1644	A	2879	109	245	QLCCFCFRQTTLIVYILSFIGMVIFTFTLDLRYI IIVFVTGGVLG
295	1645	A	2880	3	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA NNCVGEQNHRFFCALHCKSKHFCIEFTLNTNF FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS LSESISQ
296	1646	Α	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE RLQEFSQKMDQVRGHWPVST
297	1647	A	2893	8	424	SPXTLXLDTFILLGIQDNILVLILATPPFMAGG KLYSTMGRFLRDRKNPACREMAVVLLANLA QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT QIQQSQASLLHMHNPPFEPTSVDMMRRACRA LLALAKVDDNHSEF
298	1648	A	2894	310	445	FWIYFPSFFMTGYLPLGFEFAVEITYPESEGTS SGLLNASAQVNL
299	1649	A	2898	1	492	KIKAKNLTNYDLCSIFLGTSTLLVWVGVIRYL GYFQAYNVLILTMQASLPKVLRFCACAGMIY LGYTFCGWIVLGPYHDKFENLNTVAECLFSL VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD LOEF
300	1650	A	2901	1	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST TVTVRFVNKADFPKVRAKEQTFMFPENQPVS SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP FSSYEKLDITVLDVNDNAPIF
301	1651	A	2902	162	433	THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE EPCGWMYDHAKWLRTTWASSSSPNDRTFPG KPAVSEDMKELRPACSTYFNPRFPYKL
302	1652	A	2909	2	412	GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRODRAHFM
303	1653	A	2914	291	453	KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS
304	1654	A	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA
305	1655	A	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP
306	1656	A	2944	2	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA LTKGALWAVFLLAGSALLCAEVTGVIWRQPE

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	noa	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence)	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
j .				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		Ì		peptide		/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
				<u> </u>		SKTKLSFKVSSSA
307	1657	A	2950	2	411	NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL
1					1	PDLSTTEGSHAFLPCKARGSPEPNITWDKDGO
1			1			PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT
						CTAENAVGRARRRVHLTILVLPVFTTLPGDRS
L		<u> </u>				LRLGDRLWLR
308	1658	Α	2951	1	407	PTRPPRVRFDNEFDAESQRKRTTSVSKMERM
					ŀ	DSSLPEEEEDEDKEAINGSGNAENRERHSESS
1						DWMKTVPSYNQTNSSMDFRNYMMRDETLEP
i i		i		ľ	f	LPKNWEMAYTDTGMIYFIDHNTKTTTWLDP
						RLCKKAKAPEDC
309	1659	A	2954	2	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE
						LQGAVRGGAVGGSRACQRARPRGAVLG
310	1660	A	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC
						YTRDDFLFVIEHMMPLCMVISWVYSVAMTIQ
						HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI
						TGFVQLSISVTALTAILKYGQVLMHSHVVIIW
311	1661	A	2963	3	166	LFLAVYAVATIMFCF
311	1001	A	2903	3	465	MKPQMPGLGAPNGYGPGRGRAGVPGGPERR PWVPHLLPFSSPGYLGVMKAOKPGAGEGMK
						PQKPGLRGTLKPQKSGHGHENGPWPGPCNA RVAPMLLPRLPTPGVPSDKEGGWGLKSOPPS
-			•			AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPO
Į į						KI
312	1662	A	2967	3	405	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM
						HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV
1						EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS
						ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI
						DECEGPKPA
313	1663	A	2969	2	430	VVADNCRQGYLDALRFLERRGLTKEPVLWT
						LVSKEPPAPADGNWDAGCDQRRKGGLSLNW
						KVPHVQVKDVPNFEQLSPELEAALKKACTRD
j l						PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF
						RSRRLVVWLPDVPADLWWMQ
314	1664	Α	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE
[[' '					LDALGRGVFVNASGLRLLDLSSNTLRALGRH
'	[DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA
'						LSHLYLGCNELASFSFDHLHGLSATHLLTLDL
	1666		2000			SSNRM
315	1665	A.	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA
						QELYILKVMAVSGSKAELGQQTGTATVRVSI
					_	LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV
		ł	}	ŀ	,	FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ
				l		TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT
316	1666	Ā	2978	2	400	ELVVELVSAGKSGPERNTYEVQVVTGNVPKA
310 .	1000	^	2710	-	700	GTDANVYLTIYGEEYGDTGERPLKKSDKSNK
	• [ļ		FEOGOTDTFTTYAIDLGALTKIRIRHDNTGNR
		İ	l	1	ļ	AGWFLDRIDITDMNNEITYYFPCQRWLAVEE
	.	- 1		İ		DDGQLSRE
317	1667	Ā	2981	3	440	VLNCQGRPTRPVRINGDGQEVLYLAESDNVR
		·		-	. 10	LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH
		ı	- (ļ	HRENVFLSYQDKRINHGSLPHLOHRVRFAAS
			i			DPSQYDASINLMNLQVSDTATYECRVKKTTM
		1]		ļ	ATRKVIVTVQARPAVPMCWTEGQ
318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEOPKPIIILRFA
						ENYDARLLRIDIANTLREQVQELFNKTYGKQ
	ĺ	1				RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI
ــــــــــــــــــــــــــــــــــــــ	1					

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	İ	09/496	correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ		l	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1				peptide	'	/-possible nucleotide deletion, \-possible
		<u> </u>		sequence		nucleotide insertion
	1				1	STLALSHSAQVLASASGRSSTTAHCQIRVWD
	1					VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL
200	1670		2000	(02		GDHDGRTLALWGTGHL
320	1670	A	3000	693	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA
	ļ	l	ļ		•	PPFNQGFCSVYITLLNELDEAVQFSNASYEAA ILENLALGTEIVRVQAYSIDNLNOITYRFDAY
1						TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC
						GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE
1	l		1		1	GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR
<u> </u>	<u> </u>				<u> </u>	WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	A	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF
						LFPWLFLQVEVIKKAYMQGEVEFEDGENGK
222	1672	<u> </u>	2010	- 10		DGAASPRNVGHNIYILAHQLARH
323	1673	A	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH
						QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND ERVFGKRGF
324	1674	A	3020	523	797	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI
324	10/4	^	3020	323	'3' .	YLFIFFEMESHSVTHAGVORHNLNSLOPLPPG
1						FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	A	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL
	1_					GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT
	<u> </u>					FCLNKEALKDEFE
327	1677	Α	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC
						GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ
						EGDLVEVVLSASATFEDFQIRPHALTVHSYRA
	1					PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK RC
328	1678	A	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART
			3000		505	LTGALNAHNKAAVDWGWQGLIAYGCHSLV
			ŀ			VVIDSITAQTLQVLEKHKADVVKVKWAREN
	1					YHHNIGSPYCLRLASADVNGKIIVWDVAAGV
						AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI
	!					HPPNYIVLWNADTGTKLWKKSYADNILSFSF
			2000			D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED
	,					GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG
				·		RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL HRMAEKVGADITVLREREVDYDSDMPRKITE
		ł				VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL
						LGVLTQGELDNGRGRARLNLFRHLHEIQSGR
			1			TSSISFEILGFNSKGEVHGINGTQWGQTLRMG
						W
330	1680	Α	3040	3	397	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET
						LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE
		.	İ	ļ		WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ
			ŀ			VVLTMTNMGPVDTATYYCAQFARGARGSN
331	1681	A	3043	3	1600	WFDPWGQ
331	1001	^	3043	3	1509	AGIRHEAPPITSNRHRRQIDRGVTHLNISGLK MPRGIAIDWVAGNVYWTDSGRDVIEVAOMK
				ļ		GENRKTLISGMIDEPHAIVVDPLRGTMYWSD
		l		į		WGNHPKIETAAMDGTLRETLVODNIOWPTG
	1	1	. 1	l l		LAVDYHNERLYWADAKLSVIGSIRLNGTDPI
	j		1	1	•	VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV
			l	l	1	PKIHKFGHSPLVNLTGGLSHASDVVLYHQHK
				ļ	· · · · · · · · · · · · · · · · · · ·	QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG
	ł		.	- 1	ļ	KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN
				i		GGSCFLNARRQPKCRCQPRYTGDKCELDQC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	[in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]		914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		1	Ì	residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		i		peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		ļ	j			nucleotide insertion
	<u> </u>		<u> </u>	sequence		WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC
			}			TQQVCAGYCANNSTCTVNQGNQPQCRCLPG
			İ			FLGDRCQYRQCSGYCENFGTCQMAADGSRQ
			1	ſ	1	CRCTAYFEGSRCEVNKCSRCLEGACVVNKOS
			1		}	GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT
1						MNSKMMPECQCPPHMTGPRCEEHVFSQQQP
i	 					GHIASILIP
332	1682	A	3045	3	952	TTTISNFHTQVNRTYCCGTYRAGPMROISLVG
ł						AVDEEVGDYFPEFLDMLEESPFLKMTLPWGT
			į			LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD
		İ		}	1	MPKSPFKRRRSMNEIKNLQYLPRTSEPREVLF
				1		EDRTRAHADHVGQGFDWQSTAAVGVLKAV
						QFGEWSDQPRITKDVICFHAEDFTDVVQRLQ
						LDLHEPPVSQCVQWVDEAKLNQMRREGIRY
			1			ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW
]			HIRLKQYHPVVEATQNTESNSNMDCGLTGKR
			1			ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP
222	1700	<u> </u>	25.11	105	1.5	PASE
333	1683	Α	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL
						YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT
						ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG
224	1604	•	2052	27	226	CSSKTWKVAPFVRAWWRP
334	1684	A	3053	37	276	VITOLEEQLNQLTEDNAELNNQNFYLSKQLD
						EASGANDEIVQLRSEVDHLRREITEREMQLTS QKQVRRVNKVVRSLEDF
335	1685	A	3054	2	846	WDAWGDWSDCSRTCGGGASYSLRRCLTGR
333	1000	^	3034	-	040	NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA
						YNDVQYQGHYYEWLPRYNDPAAPCALKCH
}						AQGQNLVVELAPKVLDGTRCNTDSLDMCISG
			[ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC
]]						RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI
	·					TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV
	ľ					ENTTVEFORGSEROTFKIPGPLMADFIFKTRY
						TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT
			<u> </u>			CGGG
336	1686	A	3058	54	347	VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA
						LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG
						AWDWRLGSPACPHWGLHKLPRLWDPLSLYP
						VLCWGT
337	1687	Α	3059	2	709	ILTSLVELTRFETLTPRFSATVPPCWVEVQQE
j [i		QQQRRHPQHLHQQHHGDAAQHTRTWKLQT
				.		DSNSWDEHVFELVLPKACMVGHVDFKFVLN
[[ĺ			ĺ		SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF
					i	LEDHKEDILCGPVWLASGLDLSGHAGMLTLT
						SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI
		İ			ļ	CNGGMRPVVRLPSLKHQSNKGYSLASLLAK
220	1600	Ā	2000	0.5		VAAGKEKSSNVKNENTSGTRK
338	1688	Α	3060	85	384	KAFYNYHVLELLQMLVTGGVSSQLEQHLDK
						DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL
	ļ					SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE
339	1689	A	3063	236	362	CELCI SCHEMVATIEENVERREGYVAEONIVV
505	1007,	n	2003	0د2	J02 ·	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV
340	1690	A	3065	3	1249	PSSVTTMLSWV DLWQFTPLHEAASKNRVEVCSLLLSYGADPT
0.50	1050	^ .	2002	٠	1447	•
	ŀ					LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA
		,				LHCAAASPYPKRKQICELLLRKGANINEKTKE
						FLTPLHVASEKAHNDVVEVVKHEAKVNAL
ļ ļ	J		j]	j	DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN
	1					

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1 201100			***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
l					schience	
ſ	l		[peptide		/=possible nucleotide deletion, \=possible
	<u> </u>	ļ		sequence		nucleotide insertion
		j				IISLQGFTALQMGNENVQQLLQEGISLGNSEA
	i					DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE
I			[GRQSTPLHFAAGYNRVSVVEYLLQHGADVH
	ŀ		i l			AKDKGGLVPLHNACSYGHYEVAELLVKHGA
	l					VVNVADLWKFTPLHEAAAKGKYEICKLLLO
	l]			HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR
	ł					GDAALLDAAKKGCLARVKKLSSPDNVNCRD
		l				TOGRHSTPLHLAGK
341	1691	A	3070	1	547	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA
342	1071	^	3070	•	J 341	NYNYDRDSEESVINLLSYNIDGIILSEKYHTI
	l					
1		l				RTVKFLRSATIPVVELMDVQGERLDMEVGFD
1	1					NRQAAFDMVCTMLEKRVRHKILYLGSKDDT
1	ł	}		i		RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI
						HLRMQLMRDALSANPDLDGVFCTN
342	1692	A	3073	463	3	RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE
	l]			IDDNRVTAEEVDILLREGEKLAPVMAKTRILR
1				,		AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE
1						RDGLDGFITITGGKLMTYRLMAEWATDAVC
	l .					RKLGNTRPCTTADLALPGSQEPAKVP
343	1693	Α	3075	250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS
*		1			-	LGASRAQVLWFVILPGALPEILTGLRIGLGVG
i	Į.	l				WSTLVAAELIAATRGLGFM
344	1694	A	3076	2	138	LYFDAYLOSLQVAAISTFCCLLIGYPLAWAV
1 277	1094	[^	3070	2	170	
345	1695	<u> </u>	2070	460		AHSKPSTRNILLLL
343	1093	A	3078	469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC
						VINQAAATGGDARQLVGYLVSQSGLPLDTSA
						LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL
ľ	}					DRKALPLPELKAQAPGRAPKAGSETIIAAAFS
L						SLLGCDVQDADADFFALGGHSLLAMKLAT
346	1696	Α	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI
1				}		ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI
-						ALSOSRSOKELEDVLYSNLEELTRMAKMVSD
į '			ĺ			MLFLAQADNNQLIPEKKMLNLAHEVGKVFD
		}				OFEALPE
347	1697	A	3084	3	340	NELTFKEAEISKLYTKVHPAYRTLLEKROALE
1				-		DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ
1						OVYMOLLNKEOELKITEASTVGDVRIVDPAIT
240	1.000		2006	702	10	QPGVLKPKKGLIILGAI
348	1698	Α	3086	723	10	TQAMVWQQKACAEDDPQLSGRHWLHAATL
						YNIAAYPHLKGDDLAEQAQALSNRAYEEAA
]				l		QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG
]						PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI
						AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK
]						ALPNVPWVDHTRVAAFGFRFGANVAVRLAY
[[LESPRLKAVACLGPVVHTLLSGLKCQQQVPE
'				l		MYLDVLASRLGMHDASTKSSTRENH
349	1699	Α	3087	2	249	RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV
				-	-·-	GFGVAMSQALGPFSLRAGVASSTLGIAQVCG
1 1		[j	!	SSLWIWLAAVVGIGAWNM
350	1700	A	3099	3	424	EAPEATPOPSOPGPSSPISLSAEEENAEGEVSR
330	1700	Λ.	ンロフグ	,	724	
			· ·	ļ		ANTPOSDITEKTEDSSVPETPDNERKASISYFK
] l						NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN
			1			KDTVIIVSEPSEDEESQGLPTMARRNDDISELE
						DLSGMEDLK
351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG
}		ļ	í			MFGIVVMVIETELSWGAYYKAPLYSLALKCL
}		j	1			ISLFTIILLGLTIVYHAREIQLFMANYGADDWR
j l						SALTYEPIFLILLEALRGVIHATPCRVSLSLWD
		- 1	ł	ì		GLDLP
						·

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN 09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	1	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uenœ		Ī	914	ng to first amino acid	of peptide	Q=Ghttamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
1	[İ	[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł		peptide	sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
352	1702	A	3110	341	2	AQLAEVCPPQTLLTTNTSSISITAIAAEIKNPER
332	1702	^	3110	341	*	VAGLHFFNPAPVMKLVEVVSGLATAAEVVE
		l				QLCELTLSWGKQPVRCHSTPGFIVNRVARPY
	(İ		1	YSEAWRALEEQVAAPEVI
353	1703	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL
						GYGNTMVGIAVGIQFLATVLTRGYAGRLA
354	1704	A	3116	367	225	WQLFHLNGTFLNIGETDTESCVNGWYYDRSS
1		1				FPFSNMTEVRGLVFLS
355	1705	A	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF
1						EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL
1						ESRICVVGENGAGKSTMLKLLLGDL\APVRGI
						RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL
1			İ			LGTQVFLGRPEEEY\RHQLGFGMGISGELGHA
1						SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\
					'	DEPTN/HLGHGRAIEALGPCLQTISGVGVILVS
						HE*SALSRLVCRE\LWVC*GRSTSPF
356	1706	Ā	3121	137	466	RGGRDWGEHNQRLEEHQARAWQGAMDAG
						AASREHARWQGTGLAPGTRVAVAPTCVQGL
i i						PQERSVCRPFFSSRWREGPVWALGAGAHGKP
L						RWSGGVRCVVRGGRWFTPAPH
357	1707	A	3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ
1						PGLYFGGAAAVAEPDHLREAGITAVLTVDSE
j l						EPSFKAGPGVEDLWRLFVPALDKPETDLLSH
						LDRCVAFIGQARAEGRAVLVHCHAGVSRSV
1 1						AUTAFLMKTDQLPFEKAYEKLQILKPEAKMN
						EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ
						KVTEKYPELQNLPQELFAVDPTTVSQGLKDE VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH
1						KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA
]		•				LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC
1 1						GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708	Α	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP
.	' ' '	••	1 2 2 /	0.0	137	TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP
						LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP
						GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP
						PTANREINPGPAAAADTRSCWGHKRSWRGW
						RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG
						KGAGGKPSETLTRSPPVWRGKRGSANGFLSW
						VQILQ
359	1709	A	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG
'	l			ŀ		HRPLDKKREDAPNLRPALADVITVCDYRAQIA
لــــا						*AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT
1 1	.		1	Ì		AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL
						*RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL
	1-11					QA
361	1711	A	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA
			1			AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR
						VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N
	1		1	1	ł	GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA
			1			GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD
	ŀ		}		İ	PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ
		- 1	1			APGLPHRTSIRPGWRRLTEPEAWARRHRRPW GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT
	ł	i				PRLTVMSRCLAPDLKAPASGPRGWRRGMPO
'	ĺ	ſ	1	. [SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA
			.	İ		GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT
]	Į					F/LIPSPT*MSPALVIQPPVPPTQMGLRISGLPR
{	ĺ	ſ	ĺ	. [QG*PSGAPW*LPGLAQLAFQCHLPHDEVGPP
						· · · · · · · · · · · · · · · · · · ·

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	į .	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	į	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		}	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			ŀ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
<u></u>		1.	<u> </u>	sequence		nucleotide insertion .
1		1	į			RNQSPLGNDTLSSGLPMGPRRQVWPLARVG
				ł		GHSSPREPQVLKKPLWGQTDIAGVGSASLYP
	<u> </u>					DNL
362	1712	A	3136	1270	274	RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ
		1			İ	HNTWQLSRVYPSDLRTDSSNYNPQELWNAG
	ļ	İ		1		CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR
1		l		l		VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS
1		1				LKEGLEGLHGESCSSFLLGPSVAMNMQTAGL
1		ļ				EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF
1		1	1			HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE
i	ì	ľ	1		ĺ	GSIVDPLVKVQIFGVRLDTARQETNYVENNG
		l	1			FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW
]		j				KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG
-						ISLRPASIFVYICIQEGLEGDES
363	1713	C	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA
1		i -	1			PAMNESPLAPHLHQHLVFSVFQVLTILIGV**
364	1714	Α	3140	57	418	SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK
				, , ,		NHKVESPRTMSPG*DPQSFLQIPQPRPPOLRV
1			İ			GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI
]	}	1]			SGASLCPVLSPPR*PLOPSSL
365	1715	A	3145	122	413	LLPYPSLFVFLRQCHFVT\RLECNGVVSAHCN
303	1 1/13	^ ا	3143	122	413	l
	ļ	ŀ				LHLPGSSDSPASAS*VAGTTGVCHHTRLIF\VF
	ł		1			LV*TGFHYVAQAGLELLTA*S\PPQLPKVVGL
366	1716	A	2160	247	2	QA
300	1/10	A	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN
1	Ì	ļ				FIKIKKLCIEGYY/NREPQNGRKIFANYVS\DK
267	1010	<u> </u>	10.50		20.75	GLMATIYEELLKLSNKLIQ
367	1717	A	3152	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT
		ł	1			PIEKSDFAAATHPRAFYLSKPDETPNAWMSD
	ļ					SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS
	1		1			TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP
i i	İ					NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN
1 1			i i			VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE
						SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS
						ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR
						QNLKEKHARHIADLRAYYESEINSLKQKLEA
			1 1	· i		KEISGVEDWKITNQILVDRCGQLDSALHEATS
						RVRTLENKNNLLEIEVNDLRERFSAASSASKI
]			LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF
			[[ĺ		ENAYKLSDDKEAQLKQENKMFQDLLGEYES
						LGKEHRRVKDALNTTENKLLDAYTQISDLKR
						MISKLEAQVKQVEHENMLSLRHNSRIHVRPS
]		ſ	RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT
			1 1		ł	QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST
						SLLIKKORETSDTPIMRALKELDEGKIFKNWG
			j Ì		ļ	TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK
[[[j	ł	CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR
				i		EIMLTPVTVAYSPKRSPKENLSPGPSHLLSKN
]			ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV
						KKVSVRTAWEKNKSVSYEOCKPVSVTPOGN
	ľ		[[ľ		DFEYTAKIRTLAETERFFDELTKEKDOIEAAL
				}		SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718	A	3163	2	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT
300	1710	^	2103	-	0000	
1						RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL
				l	į	LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ
			-	J		MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT
				Ì	1	TVPQTQGQTTAQKVSKKTGPRCSTALATGLK
				İ		NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP
						AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLQDQLRDAQQQVKALGTERTTLEGHLAKV
					·	QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPRALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVVP MSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	A	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*1*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP ALASQSAGITGVSHLARPONLYF
372	1722	Α	3180	381	76	RVLHHDNVPAHSSPQKREISQEFQLEIRHLP*S PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL QHY*AYVEK
373	1723	A	3181	410		RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVTLLRSERPPI

SEQ ID No. of nucl- ordide pepide colide pepide colide pepide colide pepide colide pepide colide pepide colide pepide colide pepide colide pepide colide pepide colide pepide colide pep							
peptide cotide seq- uence USSN location corresponding to got the sequence provide sequence or peptide sequence or peptide sequence provide seq		, -				1	
sequence USSN glossion to design to first among to first among to first among the first among			hod	1			
uence March 1914						*****	
uence alique side state and residue of peptide sequence p							
amino acid reidised sequence peptide seq	seq-	uence	i :	1			
residue of peptide peptide sequence y=Tyrosine, X=Unknown, **Siop coden, Peptible nucleotide delicini, Peptisible nucleotide delicini, Peptisible nucleotide delicini, Peptisible nucleotide insertion FERMYDAQHQOYGSNKCRVNNAGCSSLCL ATPGSRQCACAEDQVLDADGVTCLANPSYYP PPCQCPGEPACANSRC(DERWXCDDDDCDDLCL) NSDEAPALCHQHTCPSDRTKCENNRCCTPING CASCRC[PISWTCDLDDCGDRSDESASCAYPT CPT_POTENNRCTHANDROCGINS DEAGCSISCSSTQFKCNSGRC[PEHWTCDGD NCGDYSDETHANCTON_ATPPGGCHTDEF QCRLDG;LCPILWRCDCDINDCGDNS DEAGCSISCSSTQFKCNSGRC[PEHWTCDGD NCGDYSDETHANCTON_ATPPGGCHTDEF QCRLDG;LCPILWRCDCDINDCGDNSDEKSC GTHNCYDPSVKGCGCDAGNSYWCCCSUG NNCEONSDENCESLACRPPSHPCANNTISVC LPPOKLCDGNDDCGDSDECESLACRPSHPCANNTISVC QUSGYCAHLLCSQCCDANPSYWCCSVGG WVLEPDGESCRSLDPFKPFIIPSNRHEIRRIDH KGDYSYLVFGLRNTLADPHLSQSALYWTDW VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWAGNIYWEGKCDON, SYCKCSYGG WVLEPDGESCRSLDPFKPFIIPSNRHEIRRIDH KGDYSYLVFGLRNTLADPHLSQSALYWTDW VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWAGNIYWEGKNEDYGENGLTWTDWDASLP REAAMSGAGRRTVHREITGSGGPWFONGLTV DYLEKRIL WIDDARSDAJYSARYDOSGHMEVL RHEHESGGFYGFWGGTV DAYDLARD RHENDER STANDARS AND STANDARS	uence	!	ł	914	~		
peptide sequence // possible nucleotide delizion, impossible nucleotide insertion FERMYDAQHOQVGSNKCRVNNAGCSSLCL ATRGSRQCACAEDQUADAGVTCLANPSYVP PPQCQPGEFACANSRCIQERWKCDGDMDCLD NDSAPALCHQHTCSPSWKCDGDMDCLD NDSAPALCHQHTCSPSWKCDGDMDCLD NDSAPALCHQHTCSPSWKCDGDMDCLD NDSAPALCHQHTCSPSWKCDGDMDCLD DAGGSSHCSSTDKCKSRCFERHWTCDGD DAGGSHSCSTDKCKSRCFERHWTCDGD NDCGDDYSDETHANCTNQATRPPGGCHTDSP QCRLDGCLPFLWRGCDGDTDCMSDESKSCAYPT CPPLTQPTCNNGCRIDAVRCANPPGGCHTDSP QCRLDGCLPFLWRGCDGDTDCMSDESKSC QCRLDGCDGDGDGDGGCHCDGCSLAN GCGSHRCSTAPGGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		i				of peptide	
Inaclotide insertion						sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
FERMYDAQHQVGSNKCRWNNAGCSSICL ATHGRICACAEDQUIADAGVTCLANPSYVP PPQCQPGEFACANSRCIQERWCDGDNDCLD NSDEAPALCHQHTCSPSRKCCNRRCDRW LCDGDDNCGNSEDESNATCSARTCPPNQPSC AGRCPISWTCOLD DDCQDSSDESSACSAYPT CPPLTQPTCNNGRCINWRCDNDNCGDNS DEAGCSISCSSTOKCNSGRCDEHWTCDGD NDCGDVSDETHANCTNQATRPPQGCHTDEP QCRLDGLCPPLWRCDGDTDCNDSDEKSCE GYTHYCDFSVKGCCRBARCISKAWVCDGD NDCGDNSDENGSLACRPSHPCANTSVC LPPDKLCDGNDDCGDGSDEGLCDQCSLNN GCSINCSVAPGEGICACPSPHPCANTSVC LPPDKLCDGNDDCGDGSDEGLCDQCSLNN GCSINCSVAPGEGICXPCSPHPURSWREIPERDIH KGDYSVLVPGLERNTALDFHLSQSALWYDDV VEDKYTSGKLLDNGALTSSFWOYQCLATPSG WYLEPDGSCRSLDPFPURSWRSHEIRBIH KGDYSVLVPGLERNTALDFHLSQSALWYDDV VEDKYTSGKLLDNGALTSSFWOYQCLATPSG LAYDWAGNITYWVSNLDQIEVAKLDGTLRT TILLAGDIEHFRAALDFRCGSCOWPNOLTV DYLERRI WIDARSDANTSARTYDGSGMMEVI RHEBISHFAAVILTGGGVYYTDWRTNTAR RHEBISHFAAVILTGGGVYYTDWRTNTAR RHEBISHFAAVILTGGGVYYTDWRTNTAR RHEBISHFAAVILTGGGVYYTDWRTNTAR QMARNTCRANGGCGSHLCDNYTRIYSC ACPHLMCLHEKDNTTCYERKCPLLYARQMEIR QWYYMSDWRTQARKARTHGTOYSTVVSADL PRAHGLAVDWYSRIKTWTSTDTNKKQINVA RLDGSFGNAWVQGLEQHGLVVHPLRGKLY WIDGDNISMAMMOGSRRLLFGSGGVYGFVGL ADPFESKLYWISGGNHTINRCNLDGSGLEVID ARRSQLGKATALAMBGKLWWADQVSEKM GTCSKADGSGSVVLRNSTILVMBMKVYDESI QLDHKGRTPCSVNINGSCRLVGTWEDVVTNGG RFGGRADGGGGRACA AGDFEGRANDANGSCRLFTGGGGCAACA CAHGMLAEDGASCRYJANTUNGGRACA CAHGMLAEDGASCRYVANTSTILVMBMKVYDESI QCYPREBRSRLDTGRESGCQLLTWREW GYPREBRSRLDTGRENGVVLVNVSISWNGISV DVQDGKLYWCDARTDKERGLLETGENREVV LSSNNMDMFSVSVERGDFTYWSDGTHANGSIK RGSKDNATDSVPLRTGIGGVQLKDRVFNRDR QCYPREBRSRLDTGRENGVVLVNVSISWNGISV DVQDGKLYWCDARTDKERGLLETGENREVV LSSNNMDMFSVSVERGDFTYMSDRTHANGSIK RGSKDNATDSVPLRTGIGGVQLKDRVFNRDR QCYPREBRSRLDTGRENGVVLVNVSISWNGISV DVQDGKLYWCDARTDKERGULETGENREVV LSSNNMDMFSVSVERGHYDFDLANVTTITIT RHTVDQTRCGARERETVITMSGDDHFRAFVL LIEKDRTPGJANDGGRACA CAHGMLAEDGASCRSYAGYLLYSERTLLKSI HSDERNIAMPGVGFPEDFHMKNNTALAFDV RAGGSEDHANDFFFINGANGDLETGENREVV LENNMDMFSVSVERGHPTYMSDGTHANGSIK RGSKDNATDSVPLRTGIGGVQLKDRVFNRDR QCGTPLDFRANKLYCHANGSCCHARCA CAHGMLAEDGASCRSYAGYLLYSERTLLKSI HSDERNIAMPGVGFPEDFHMKNNTALAFDVT RHTVDQTRCGARERETVITMSGDDHFRAFVL LIEKDRTPGGANGCCHAACH CHARCHTON CHARCH RTTTART	j	ł	ł		peptide		/=possible nucleotide deletion, \=possible
ATTGSRÇCAĞA EDQVILADAGYTCLANPSYVP POCOPGEACANSR (CIGRWYCCDIDNOCL) NSDEAPALCHOHTCPSDRIKCENNRCIPNRW LCDGDIDCGNSEDES ANTCSARTCPROYSC ASGRCPISWTCDLDDCGDRSDES ASCAYPT CPITOTTCNINGCRINIWRCDNIDOCGDNS DEAGCSISCSSTORKONSCRCIPEIWTCDGD NCGDYSDETHANCTNOANTSPOOGCHTDEP QCRLDGLCPIRWRCDGDTCMDSDEKSCE GYTHYCDFSVKFGCKDSAKCISKAWVCDGD NCEDNSDENCESLACRFSHCANTSVC LPPDKLDGADDCGDGSDEGIC COCCSLINI GCSSINGCSAPGAGTCSCCLGMELLEPDNHT CQIGSYCAKHLKCSQKCDQNKFSVKCSCYEG WYLEPDGESCSLDPFKPFIEINSREIRRIDLH KGDYSVLVPGLRATIALDFHLSQSALYWTDV VEDKIYKGKLLDNGALTSSEVVQVGLATFEG LAVDWAGNIYWVESNLDQIEVAKLDGTLRT ILLAGDIEPPRALADFROELFFYDDVDASLP REAASMSGAGRRTVHEETGSGWPNGLTV DYLEKRILWDARSDAYTSARYDGSGHMEVL RCHEFLSHFFAVTLYGGEVYWTDWRTNTLAL KANKWTGNVTVVGNTNOFPDLGVYHPS QPMAPNCEANGQOPCSHLCLNYNRTVSC ACPHIRMCHEMONATORYTUSPT (VICTOR) GWADNINGCRINITYSTYDDDNVTVLDYDARE GVDLDAPYYNYISFTVDDDNVTVLDYDARE GVDLDAPYYNYISFTVDDDNVTVLDYDARE GVDLDAPYYNYISFTVDDDNVTVLDYDARE GVDLDAPYYNYISFTVDDDNVTVLDYDARE GVDLDAPYYNYISFTVDDDNVTVLDYDARE GVDLDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GVDLDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNISFTVDDDNVTVLDYDARE GROUDAPYYNYISFTVDDDNVTVLDYDARE GROUDAPYYNISFTVDDDNVTVLDYDARE GROUDAPYYNISFTVDDDNVTVLDYDARE GROUDAPYYNISFTVDDDNVTVLDYDARE GROUDAPYYNISFTVDDDNVTVLDYDARE GROUDAPYYNISFTVDDNVTVLDYNTSGCLOVD AMRSOLGKATALAMGDSLWWADOVSEKM GROESKANDART GROUDAPH GROUDAPH GROUDAPH ARDD GROUDAPH	L	<u> </u>		ľ	sequence		nucleotide insertion
PPQCQPGEFACANSRCIGERWICDGINDLCLD NDSEAPALCHQHITCPSDRIKCENRICEINRW LCDGINDCGNISEDESINATCSARTCPPROPSIC ASGRCIPISWTCOLD DDCGDRISESSACAYPT CPPLTQPTCNINGRCHINWRCDNDNCGDINS DEAGCSISCSSTORKOSGRCIPERWITCDGD NDCGDYSDETHANCTINQATRPPGGCHTDEF QCLDGLUPILWRUCDGYDTCMDSSDEKSCE GVTHYCDPSVKFGCKDSARCISKAW VCDGO NDCGDVSDETHANCTINQATRPPGGCHTDEF QCLDGLUPILWRUCDGYDTCMDSSDEKSCE GVTHYCDPSVKFGCKDSARCISKAW VCDGO NDCGDNISEBECISLACRPSHICANNTSVC LPPRKLCDGNDDCGGBDGEGLCGCSLIN GGCSHINGSYAGEGIVCSCFLGMELGPINHT CQIQSYCAKHLKSQKCDQNKFSYKKSCYEG WVLEPDGESCRSLDPFKFFIEISNRHEIRRIDLH KGDYSVLVFGLRNTLADEHLSGSALYWTDV VEDKIYRGKLLDNGALTSFEVVQVGLATPEG LAVDWAGNTYWVESNLDQIEVAKLYWTDV VEDKIYRGKLLDNGALTSFEVVQVGLATPEG LAVDWAGNTYWVSRNDQIEVAKLOTTLRT TILLAGDEIPPRAIALDPROGLFWTDWDASLP REASANGSGRRTVHERTGSGGWPNGLTV DYLEKRILWIDARSDATSSARYDGSGHMEVL RHESHFFAVILYGGGVYWTDWRNSLY RHESHFFAVILYGGGVYWTDWRNTILA KANKWTGINVTVVQRINTOPPDLQVYHESR QMAMPICEANGGQFCSHLCLINYRITVSC ACPHLMGLHKDNITCYEFKKFLLYARQMEIR GVDLDAYYSYNISFTVDDNAVTLDYDARE QRYWSDWRTOAKKAFBIGTOVETVVSADL PAAHGLADWYSSINFWTSYDDTNKKQINVA RLDGSFRNAVVQGLEQPHGLVVFPLRGKLY WTDGDNISMANDGSRTLLFSGQKGPVGI ADPFESKLY WISSGRHTINRGNLDGSGLEVID AMRSQLGKATALAIMGDKI.WWADQVSEKM GTSKAADGSGSVIRASTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVYNGG RKPGLADWINSTILVPBKGKLYVYNGG RVGLADWINSTILVPBKGKLYVTYNGG RVGLADWINSTILVPBKGKLYVTYNGG RVGLADWINSTILVPBKGKLYVTYNGG RVGLADWINSTILVPBKGKLYVTYNGG RKGTNVCAVANGGCQQLCLTYRGRGGRACA CAHGMLADGASGRYAGGGVGSFLLLVFWTEW GQYPRIERSRLDGTERVVLVNGSWPRGIS DYDGKATALAGGGRACA CAHGMLADGASGRAFAGYNGTUNFBCRULCTGRINEVV LISSNNMMFSVSVFEDFIVWSDRTHANGSIK RSKDNATDSVLRTGIGGQULGDWFRDR RGSTDVYNGGLAPRERETVTINSGDDHPRAFVL DECNIMAPONTNWEQGFBGRRALAGANVL TILBKDRTNGLADWINGVRINGGLALCRUNGSCE RCEYDGSHRYVLKSFRYFGGLAVYGGHF WTDWVRRAVQRANKHVGSNMKLLKVDIPQ QPMGILAVANDTNSCELSEPCRNNOGCQDLCL LTHQGHVNCSCGGGRLQDDLTCANNSSCE AGDDEGGERAGCCANAGCCOS AGDDCGGGSGSGLCDDLTCANNSSCE AGDDCGGGGSGSGSGLA							
NSDEAPALCHOHTCFSDERKCENNRCIPTMY LCDGDNDCONSEDESNATCSATCPPNOPSC ASGRCIPISWTCDLDDDCGDRSDESASCATY TYPY TOPICTONING CININWRCDDDNDCGDNS DEAGCSHSCSSTOPKCNSGRCIPEHWTCDGD NDCGDVSDETHANCTODG NDCGDVSDETHANCTONGTPROGCHTDER QCRLDGLCFLRWRCDGDTDCMDSSDEKSCE GVTHVCDPSVLFGCKDSARCISKAWCDGG NDCEDNSDEENCESLACRPSHPCANNTSVC LPDRLCDGNDDCGDGSDEGELOQCSLNW LGCDNSDEENCESLACRPSHPCANNTSVC LPDRLCDGNDDCGDGSDEGELOQCSLNW LGCSHNCSVAPGGIVCSCPLGMELOPDNIV CQUSYCAKHLKCSQKCDQNRFSVKCSCYPG WVLEPDGESCRSLDPFRFHIFSNRHERRIDLH KGDYSVLYPGLRYTTALDFHLSQSALYWTDV VEDKIYRGKLLDNGALTSFEVVIYGGLATFEG LAVDWAGNTYWESNLDQIEVAKLDGTLAT TILAGDEHPRALDPRDGLEPVTUVMDASLP RIEAASMSGAGRATVFRETGSGGWPNGLTV DYLEKRIL WIDASADATYSARYDSGGHMEVV RGHEFLSHPFAVTLYOGEVYWTDWATSTT AKANKWTGHNYTVQRTNYCPPDLQVYHPSR QPMAPPCEANGGGGPCSHLCLINYNRTVSC ACPHLMCLHKDNTTCYFFKKFLLYARGMER GVDLDAPYNYTISTTVPDDNVTVLDYDAGA RLDGSFRNAVVQGLEQPHGLVWHPLRGKLY WTDGDNISMANMDGSNRTLLFSGGKGPVGL AMPRICALAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	İ	ļ					ATPGSRQCACAEDQVLDADGVTCLANPSYVP
LCDGNNDCONSEDSANTSARTCPPNOPS ASGRCPISWTCDLDDDCGRSDSASCAYPT CPFLTOFTCNNGRCININWRCDNDDCGDDN DGAGCSHSCSSTOPK CNGRCIPEWTCDGD NDCGDYSDETHANCTNOATRPGGCHTDEF QCRLDOLCPLRWRCDGDTDCMSDEKSCE GVTHVCDPSVKFGCKDSARCISKAWVCDGD NDCEDNSDENCESLACRPSHFCANTTSVV LPDREACHOLOGGDEGSPEGELCQCSLNN GGCSHNCSVAPGEGIVCSCLGMELGPDNHT CQIGSYCAKELKCSQKCDNKFSVKCSCYEG WVLEPDGESCRSLDPFKPFIFISNRIERRIDH KGDYSVLYPGENTALDFHLSSALYWTDV VEDKIYRGKLLDMGALTSFEVVQQGLATFEG ELAVDWAGNTYWYSSNLDQIEVAGNTAL FRIAADELFRALDFRDGLFWTDWDASL PRIAASMSGAGRTVYRETGSGGWPNGLTV DYLEKRIL WIDARSDAYSARYDGSGHMEV, RGHEFLSFFAVLYGGGVAWTUNDYSALV RGHEFLSFFAVLYGGGVAWTUNDYASLE PRIAASMSGAGRTVYRETGSGWPNGLTV DYLEKRIL WIDARSDAYSARYDGSGHMEV, RGHEFLSFFAVLYGGGVAWTUNDYASLE QRAYPYGCANGGQGCSGLCLANYRTVSC ACPHLMKLHKDNTTCYFFKKFLLANRAWGA GAPAPPCEANGGGGCSGLCLANYRTVSC ACPHLMKLHKDNTTCYFFKKFLLANRAWGA RCHAMKAWTGHNTVVQRTNTQFFDLQVYHFS QRAYWSDVRTQAKRARINGTGVETVVSADL PNAHGLAYDWYSRALPSTSYDTINKCQDWA RLDGSFNAAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGVVGL ADPPSSKLYWISSGNITINKCHLOSGLEVID AMRSQLGKATALAMGGSNLWADQVSEKW GTCSKADGSGSVLYRNSTTLVMHKVYDESI QLDHKGTNFCSVNNGDCSQLCLPTSSTTRSC MCTAGYSLRSGQQACGSCNGSTLYSVHEGIG GPILDPHOMSDALVPNSGTSLAVGDPHAEDD TYDVOMGLSTISKARRAFOTWEREVYTNGIG RVEGIAVDWIAGNITWTTOGGTDVETVEND SRYVYNSQLOKPRATIVHTENGULFWTHE GQYPRIERSRLDGTGRVVVNNSISWYHGIS SRYVYNSQLOKPRATIVHTENGULFWTHE GQYPRIERSRLDGTGRVVVNNSISWYHGIS SRYVYNSQLOKPRATIVHTENGULFGRNRVV LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATDSVPLRTGTIGVQLOKPTENE GQYPRIERSRLDGTGRVVVNNSISWYHGIS SRYVYNSQLOKPRATIVHTENGULFGRNRVV LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATDSVPLRTGTIGVQLOKPTENE GQYPRIERSRLDGTGRVVVNNSISWYHGIS SRYVYNSQLOKPRATIVHTENGULFGRNRVV LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATDSVPLRTGTIGVQLOKACAGGGGGLUFTSTTTT RHTVDQTRPGAFERTVITMSGDDHPRAFV LDECQNLMFWTWNFQHTSIMAALGGANVL TLIERDRTFNGLADHARKLTYTTT RHTVDQTRPGAFERTVITMSGDDHPRAFV DDCQNLMFWTWNFQHTSIMAALGGGQDLC LTHQGHVNCSCRGGRILLQDDLTCRAVNSSCE ACDEFECANGGCNSLUCVSMALL		1					PPQCQPGEFACANSRCIQERWKCDGDNDCLD
ASGRCIPISWTCDLDDDCGDRSDESASCAYU CFPLTOFTCNIGGRINNWRCDDDNDCGDNS DEAGCSHSCSSTOFKCNSGRCIPEHWTCDGD NDCGDVSDETHANCTNOATRPGGCHTDEP QCRLOGLCTLRWRCDGDTDCMDSSDEKSEE GYTHVCDPSVKFGCKDSARCISKAWVCDGD NDCEDNSDEERCESLACEPPSHFCANTSVC LPPDRLCDGNDCGDGSDGGELCDQCSLNN GGCSHCNSVAPGEGIVCSCTLGMELGPDNHT CQIGSYCAKHLKCSQKCDQNKFSVKCSCYEE WVLEPDGSSCSSLDPFKFFIFISRHERRIDLH KGDYSVLVPGLRITIALDFHLSQSALYWTDV VEDKIYRGKLLDMALTSFEVVLYQGLATTPG LAVDWLAGNTYWVESNLDQIEVAKLDGTLTT TLLAGDEHPRALADPROILFYTDWDASLP RIEAASMSGAGRRTV-IRETGSGGWPNCLIV DYLEKRIL WIDASADLYSARTYDGGGHMEVL RGHEFLSIFFAVTLYQGEVYWTDWRTNTLA KANKWTGHNYTVQRTNTQPTGYCYHTWSC ACPHLMKLERDNTTCYEFKKFLLVARQMEE GWYWSDVRTQALKRAFINGTGVETVVADAL RJCSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDJNSMAMNDGSNRLLFSTYDTINKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYDTINKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYDTINKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYDTINKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYDTINKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYDTINKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYTDYNKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYTDYNKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYTDYNKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYTDYNKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKLY, WTDGDNISMAMNDGSNRLLFSTYTDYNKQINNA RLDGSFKNAWVQGLCQPHGLVVHPLRGKCY, AMPSQLGKATALAMMOKLWWADQVSEKM GTCSKADGSGSVLRNSTTLWHMKSVYDEBOHAND TYWVDMGLSTISRAKRDQTWREDVYTNGIG RVEGIAVDWLAGNTYWTDQGFDYEVEVARM GTCSKADGSGSVLRNSTTLWHMKSVYDEDGFAND TYPWVDMGLSTISRAKRDQTWREDVYTNGIG RVEGIAVDWLAGNTYWTDQGFDYEVEVARM GROTNYCAVANGGCQLCLTYRGRQCACA CALGMALEDGASGREVAGVTLYSRTTISTIT RHTVOTORTPGAFFRTTVITNSGDHFGRAVL LSSNNMDMFSVSYFEDFTYNSDRTHANGSIL RCEYDGSHRYVLKSERVHPGLAVYGGHF WTDWVRRAVQRARTHKERGDLTGGCQDLC LTHQGHYNCSCRGGRILLQDDLTCRAVNSSC AQDEFECANGEGNESLTCCONYTHCKOKSDE KPSYCNSRRCKKTTFQCSNGRCVSNMLLKVDIP		1					NSDEAPALCHQHTCPSDRFKCENNRCIPNRW
CFFLTOFTCNNGECNINNWCDNDNDCGDN DEAGCHSCSSTOFK CNGSGCIPEHWTCDGD NDCGDYSDETHANCTNOATRPGGCHTDEF QCRLOALCIPLE WRCDGDTDCMSDESSCE GVTHYCDFSVCFGCKDSARCISKAWVCDGD NDCEDNISDESCESLACRPSHPCANTTSVC LPPDKLCDGNDDCDDGSDEGELCDQCSLNN GGCSINCSVAPGEGIVCSCTLGMELIGPDNHT CQGSYCAKHLKCSQKCDQNCFSVKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKCSCVEG WVLEPDGESCRSLDFFKFHIFSNRHERRIDLH KGDYSVLYPGLRYTALDFHLSSYKLDFTLAT TLLAGDEHFRALDDFALLFYTDWDASTLY UPLEKRLL WIDARSDAIYSARYDGSGHNEILY DYLEKRLL WIDARSDAIYSARYDGSGHNEILY RGHEFLSFFRAVLY GGEVYWTDWRTNTLA KANKWTIGHNYTVVQRTNTGPFDLQYVHFSS QMANFYFCEANGGGPCSILCLINYNRTVSC ACPHLMKLHKDNTTCYFFKKFLLYARQMER GVALDAPYYNTISTTYDDMOTYDARE QRYYWSDVRTQALRARANGTGVETVVSAC ACPHLMKLHKDNTTCYFFKKFLLYARQMER GVALDAPYYNTISTTYDPDDNYTUNKQDNVA RLDGSFRAVVQGLEQPHCLVYHPLRCKLY WTDCDNISMANNDGSNRTLLSOKGEV/GL ALDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGRATALAMGGNKLWADQVSEKKM GTCSKADGSGSVYLNRSTILVMHMKYYDESI QLDHRGTHPGSVANGDCSQLCLPTSETTRSC MCTAGYSLRSGQACESWOSTLYSYHEGIG GPLDPHDIKSDALVPNSGTSLAVGIDPHAEDD TYVWDMGLSTSRAKRDGTWREDVYTNGIG RVEGIAVDWIAGNIYWTDGGFDVEVARING SFRYVYNSQLDKPAATIVHERGVJLFWHEW GQYPRIERSRLDGTERVVLNVSISWYNGIS DYQDGKLYWCDAATTNKERBULTGGRNEVV LSSNNMDMFSVSYFEDFTYWSDRTHANGSIK RGSKDNATDSYPLRTGIGIQVLKDYSRTILKSI H.SDERNLNAPVOPFEDPEIMKNNALAFDY RACTSFDUFFRHIFSDHEFGINGDAGCRAT VENNGSGELAVHROWDTLYWTSSTTTT RHTYDQTRPGAFERETVITMSGDDHPAATV TUENNGSWEGLAVHROWDTLYWTSSTTSTT RHTYDQTRPGAFERETVITMSGDDHPAATV TLLEKDIRTPNGLADHAERKLYFSDHFLING QCYNKENSTLORGCQDLCL LTHQGHYNCSCRGGRILLQDDLTCRAVNSSCE ACDEFFCANGEGNTSLTCDGVYFKCNKSDE KYSYCNSRRCKKTFRQCSNGRCVSNMLLWCDI	•	ł					LCDGDNDCGNSEDESNATCSARTCPPNQFSC
DEAGCSHSCSTORKCNSGRCIPEHWITCDGD NDCGDYSDETHANCTNOATPGGCHTDEP QCRLDGLCIPLAWCROGDTDCMSSDESSES GVTHVCDPSVKGGCKDSARCISKAWVCDGD NDCEDNSDEEMCESLACRPPSHRCANTTSVC LPPDKLCDGNDDCDGDSDEGELCDQCSLNN GGCSHNCSVAPGEGIVCSCPLOMELGPDNHT CQIQSYCAKHLKCSQKCDQWKFSVKCSCYEG WVLEPDGESCRSLDPFKPFHIPSNHEHRIBLIH KGDVSYLVFGLRNTIALDHLSQSALVWTDV VEDKYRGKLLDNGALTSPEVVIQYGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TLLAGDEHPRALDAPRAGETWTDWDASLP RIEASMSGAGRRTVHRETGSGGWPGLIV DYLEKRILWIDARSDATYSARTDSGGMEVL RGHEFI.SHEFAVILYGGGYVWTDWRTNILA KANK WTGHNVTVVQRTNTOPPDLQVYHESR QPMAPNCEANGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYERKFLLLYARQMER GVDLDAPYTNYISFTVPDDNVTVLDYNADL RGHEFI.SHEFAVILYGGGYVGL WTDDATSMAMMDGSNRTLLYSQKGYGGL PNAHGLAVDWYSRNLFWSYSTYNKKQINVA RLDGSFKNAVVGGLEPHGLVVHFRGKLY WTDDATSMAMMDGSNRTLLTSGGKGPVGL ALDFESKLYWISSCNHTNRCNLOSGLEVID AMRSQLGKATALAMIODRIK WADAQVSEKW GTCSKAADGSGSVVLRNSTILVMHMKYVDESI QLDHKGTINCSVNNGGCSQLCLTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GPLDPINKSDALPYSGTSLAUPDHAND TIYWVDMGLSTISRAKRDQTWREDVYTNGIG RVGRIAVDWAGNLYWTGGFBVEYALLANG SFRYVVISQGLDKPRAITVHPEKGYLPWTEW GQYPRIERSRLDGTBEXVLVNNSISWTHOISV DYQDGKLYWCDARTDKERDLETGENREVV LSSNNAMDMFSVSPEEDTYWSDRTHANGSIK RGSKDNATIDSVILRIGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGGRACA CAHGMLAEDGASCRYVGNLYSRTTILYSRTTIST HTWPOGNSTRAGETBYNLYSTITT RHTVDQTRTGAGGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST HLDGRGRACACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTIST RHTVDQTRTGAGAGRACA CAHGMLAEDGASCRYTAGTLYSRTTICH RTTDQTRTGAGAGRACACA CAHGMLAEDGASCRYTAGTLYSRTTICH RTTDQTRTGAGAGRACACACHORSLCTLOXGONNALLATDFY RAGTSGTDTRNTTSTTT RHTVDQTRTGAGAGRACACACHORSLCTLOXGONNALLATDFY RAGTSGTDTRNTTSTTSTTTTTT RHTVDQTRTGAGAGRICADDALACTCANTOSCR AQDEFECANGGCGCHCSLTCLOXGONNALLATDFY RAGTSGTDRN							ASGRCIPISWTCDLDDDCGDRSDESASCAYPT
NDCGDYSDETHANCTNOATRPEGCHTDEP QCRLDGLCFURWRCDGDTDCMDSSDEKSCE GVTHVCDPSVKGCKDSARCISKAWVCDGD NDCEDNSDEENCSLACRPEANTSVC LPPDKLCDGNDDCGDGSDEGELCDQCSLNN GGCSHNCSVAPGGIVCSCPLOMELGPDNHT CQQSYCAKHLKCSQKCDQWKSVKCSCYEG WVLEPDGESCRSLDPFKPFIIFSNRHEIRDLH KGDVSVLVFGLANTALDFHLSQSALVWTDV VEKYKTGKLLDNGALTSFEVVIQYGLATFG LAVDWIAGNIYWESNLDQJEVAKLDGTLRT TLLAGDIEHPRAIALDPROGLEWTDVASLP RIEAASMSGAGRTVHRETGSGWPNGLTV DYLEKRILWIDARSDAYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGGYGFYGLATFG AKANKGRRTVHRYDTDWRTNTLA KANKWTGHNVTVVQRTNTQPFDLQVYHPSR QPMAPNFCEANGGGFCSHLCLINNTRYSC ACPHLMGLHENTTCYEFKKFLLYARQMEIR GVDLDAPYYNJISFTVPDIDNVTVLDYDARE GVPLAPAYTYGGFCHYVSADL PNAHGLAVDWVSRNLFWTSYDTINKQNVA RLDGSKNAWVQGLEQPHGLVYHPLRGKLY WTDGDNISMANMOGSRTILFSGGKGPVGL ALDFFESKLYWISGSMCHTNCLOSGLEVUD AMRSQLGKATALAMODKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVHGRGKGPVGL ALDFFESKLYWISGSHCHNCSGCEVUD AMRSQLGKATALAMODKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVHGRGKGPVGL AMPFESKLYWISGSHCHYNCTORGY GPHADALDPHOKSDALVPGGTSLAVGUPHAEND TIYWVDMGLSTISRAKRDQTWREDVYNGIG GPLDPHOKSDALVPGGTSLAVGUPHAEND TIYWVDMGLSTISRAKRDQTWREDVYNGIG GPLDPHOKSDALVPGGTSLAVGUPHAEND TIYWVDMGLSTISRAKRDQTWREDVYNGIG GPRESRLAVGHAGNIYWTDGGFDVEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLAGTGRAVLNYSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNADMFSVSYEDETFWNGQNDDGSRRT TYENGGSVGLAYFRGWDTLAYSGTTLING RGSKDNATDSVPLRTGIGVOLKDIKVFNRDR QKGTNVCAVANGGCQLCLYRGRGQRACA CAHGMLAEDGASCRYAGYLLYSERTILKSI HLSDERNINAPVQFFEDPEHMKNIVLALFDY RAGTSGCTPRISTSDHFGNINQQNDDGSRRT TVENVGSVGLAYFRGWDTLYWTSYTTSTT RHTVDQTRFGAACRSCFYGTLYSETTLINST THENDGTRFGAACRSCFYGTLYSETTLINST THENDGTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTC ACCERCIONACCORD ACCERCIO							CFPLTQFTCNNGRCINENWRCDNDNDCGDNS
NDCGDYSDETHANCTNOATRPEGCHTDEP QCRLDGLCFURWRCDGDTDCMDSSDEKSCE GVTHVCDPSVKGCKDSARCISKAWVCDGD NDCEDNSDEENCSLACRPEANTSVC LPPDKLCDGNDDCGDGSDEGELCDQCSLNN GGCSHNCSVAPGGIVCSCPLOMELGPDNHT CQQSYCAKHLKCSQKCDQWKSVKCSCYEG WVLEPDGESCRSLDPFKPFIIFSNRHEIRDLH KGDVSVLVFGLANTALDFHLSQSALVWTDV VEKYKTGKLLDNGALTSFEVVIQYGLATFG LAVDWIAGNIYWESNLDQJEVAKLDGTLRT TLLAGDIEHPRAIALDPROGLEWTDVASLP RIEAASMSGAGRTVHRETGSGWPNGLTV DYLEKRILWIDARSDAYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGGYGFYGLATFG AKANKGRRTVHRYDTDWRTNTLA KANKWTGHNVTVVQRTNTQPFDLQVYHPSR QPMAPNFCEANGGGFCSHLCLINNTRYSC ACPHLMGLHENTTCYEFKKFLLYARQMEIR GVDLDAPYYNJISFTVPDIDNVTVLDYDARE GVPLAPAYTYGGFCHYVSADL PNAHGLAVDWVSRNLFWTSYDTINKQNVA RLDGSKNAWVQGLEQPHGLVYHPLRGKLY WTDGDNISMANMOGSRTILFSGGKGPVGL ALDFFESKLYWISGSMCHTNCLOSGLEVUD AMRSQLGKATALAMODKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVHGRGKGPVGL ALDFFESKLYWISGSHCHNCSGCEVUD AMRSQLGKATALAMODKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVHGRGKGPVGL AMPFESKLYWISGSHCHYNCTORGY GPHADALDPHOKSDALVPGGTSLAVGUPHAEND TIYWVDMGLSTISRAKRDQTWREDVYNGIG GPLDPHOKSDALVPGGTSLAVGUPHAEND TIYWVDMGLSTISRAKRDQTWREDVYNGIG GPLDPHOKSDALVPGGTSLAVGUPHAEND TIYWVDMGLSTISRAKRDQTWREDVYNGIG GPRESRLAVGHAGNIYWTDGGFDVEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLAGTGRAVLNYSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNADMFSVSYEDETFWNGQNDDGSRRT TYENGGSVGLAYFRGWDTLAYSGTTLING RGSKDNATDSVPLRTGIGVOLKDIKVFNRDR QKGTNVCAVANGGCQLCLYRGRGQRACA CAHGMLAEDGASCRYAGYLLYSERTILKSI HLSDERNINAPVQFFEDPEHMKNIVLALFDY RAGTSGCTPRISTSDHFGNINQQNDDGSRRT TVENVGSVGLAYFRGWDTLYWTSYTTSTT RHTVDQTRFGAACRSCFYGTLYSETTLINST THENDGTRFGAACRSCFYGTLYSETTLINST THENDGTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTT THUTDQTRFGAACREPATTYTSTC ACCERCIONACCORD ACCERCIO							DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD
GVTHVCDPSVKGGCNSARCISKAWVCDGD NDCEDNSDEROESLACRSPEPCANTSVC LPPDKLCDGNDDCGDGSDEGELCDQCSLNN GGCSHNCSVAPGGIVCSCPLGMELGPDNHT CQIQSYCAKHLKCSQKCDQWKSVKCSCYEG WVLEPDGESCRSLDPFKPFIIFSNRHEIRDLH KGDVSVLVFIGLNRTIALDHFLSGSALVWTDV VEKYYRGKLLDNGALTSFEVVIQYGLATPG LAVDWIAGNIYWESNLDQIEVAKLDGTLRT TLLAGDIEHPRAIALDPRDGLLFWTDWASLP RIEAASMSGAGRRYVHRETGSGGWPNGLTV DYLEKRILWIDARSDATYSARYDGSGHMEVL RGHEFI.SHPFAVILYGGGYWGLTV RGHEFI.SHPFAVILYGGGYWGLTV DYLEKRILWIDARSDATYSARYDGSGHMEVL RGHEFI.SHPFAVILYGGGYWGLTV RGHEFI.SHPFAVILYGGGYWGLTVYGARL RGHEFI.SHPFAVILYGGGYWGTWYGYTDWASLP REPAASMSGAGRYVHRETGGGGWPNGLTV DYLEKRILWIDARSDATYSARYDGSGHMEVL RGHEFI.SHPFAVILYGGGYBLCLINYNTYCSC ACPHLMGLHKDNITCYEFKFLLYARQMER QVDLDAPYYYSISFIVPDIDNVTVLDYDRE GVPLDAPYYGSDRITGYFIVPDIDNVTVLDYDARE QRVYWSDVRTQAKRAFINGTGYETVVSADL PNAHGLAVDWVSRNLFWTSYDTINKGNWA RLDGSFKNAVVQGLEQPHGLVYHFLRGKLY WTDGDNISMANMOSSRTILLFSGGKGPVGL AUDFESKLYWYSDSTILLFSGGKGPVGL AUDFESKLYWYSSGNHTHRGUVHFLRGKLY WTDGDNISMANMOSSRTILLFSGGKGPVGL AUDFESKLYWYSSGNHTHRAGGSGLEVD AMRSQLGKATALAMODKLWWADQVSEM GTCSKADGSGSVVLRNSTILLFSGGKGFVGL AUDFESKLYWISSGNHTHRAGWYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQQACEGVGSTLLYSVHGGR GIPLDPINKSDALVPYGGTSLAVGUPHAEND TIYWVDMGLSTISRAKRDQTWREDVYNGIG RVEGIAVDWLAGNITYWTDGFDVIEVARLNG GFYYVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSLAGGTEVYAFUTBW GQYPRIERSLAGGTEVYAFUTBW GQYPRIERSLAGGTEVYAFUTBW GQYPRIERSLAGGTEVYAFUTSTITT HYWYGACAGACQCQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNINAPVQFFEDPEHMKNVILAFDY RAGTSGCTPRIGTSDHFAGNIQQNDDGSRRT VENGSVEGLAYFRGWDTLYWTSYTTSTIT RHTYDQTRFGAFERSTVITROQQNDDGSRRT VENGSVEGLAYFRGWDTLYWTSYTTSTIT RHTYDQTRFGAFERSTVITROQQNDDGSRRT VENGSVEGLAYFRGWDTLYWTSYTTSTIT RHTYDQTRFGAFERSTVITROGCODLICL TIEGRIFYNGSCGGRILQDDLTCRAVYNSCR AQDEFECANGGGGRILQDDLTCRAVYNSCR AQDEFECANGGGGRILQDDLTCRAVYNSCR AQDEFECANGGGGRILQDDLTCRAVNSCR AQDEFECANGGGGRINSLTCDGVFRCNKNSCR AQDEFECANGGGGRINSLTCDGVFRCNKNSCR AQDEFECANGGGGRINSLTCDGVFRCNKNSCR							
NDCEDNSDEMCESLACRPSSIPECANNTSVC LPPPKL CODRDDCOGDSEDELCDQCSLNN GGCSHNCSV APGEGIVCSCPLGMELGPDNHT CQIQSY CAKELKCSQKCDQNKFSYKCSCYEG WYLEPGESCRSLDPFKPIIFSNNEBERRIDLH KGDYSVLVPGLRNTALDFHLSQSALYWTDV VEDKIYRGKLDNGALTSEVIQYGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TILLAGDEIPFRALLDPROGLFWTDWDASLP RIEAASMSGAGRRTVIRETGSGGWPNGLTV DYLEKRI WIDARSDAIYSARYDGSGHMEVL ROHEFISHPFAVILTGGEVYWTDWRTNTLA KANK WTGHNYTVVQRTNTQFDLQYYHFSR QPMAPNPCEANGGGOPCEANGGGOFCH AVARYMSDVRTQAIRAFTGGFTVVSADL PNAHGLAVDWSRNLFYSTDTNKKQDNS QRYYWSDVRTQAIRAFTNGTGVETVVSADL PNAHGLAVDWSRNLFYSTDTNKKQDNS RIEGSKNAVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSOQKGPVGL AIDFPESKLYWISGNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACGFSLLYSVHGGIR GFLDPNDKSDALVPVSGTSLAVGIPFHAEND TITWVDMGLSTISRAKROTWREDVYTNGIG GFLDPNDKSDALVPVSGTSLAVGIPFHAEND TITWVDMGLSTISRAKROTWREDVYTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARING SFRYVVISQCILDRFAITVHPEKGYLFWTEW GQYPREBRSLDGTERVVLNNSISWPNGISV DYQDGKLYWCDARTDKERDLCTEENREVV LSSNNMDMFSSVFLEFTYWSDRTHANGSIK RGSKDNATDSVPLRTGIGGQULKDIKVPNRDR QKGTNVCAVANGGCQULZPRGRORACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLAPVOFFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGIQQNDDGSRRT THENVGSVEGLAYHRGWDILYWTYTSTTSTT RHTVDQTRFGAFFRETVITMSGDDHPRAFVL DECQNLMFWTNFRFSDIHFFGIQQNDDGSRRT TIVENVGSVEGLAYHRGWDILYWTYTSTTSTT RHTVDGTFGAFFRETVITMSGDDHPRAFVL DECQNLMFWTNFRIFFSDIHFFGIQQNDDGSRRT TIVENVGSVEGLAYHRGWDILYWTYTSTTSTT RHTVDGTFGAFFRETVITMSGDDHPRAFVL DECQNLMFWTNFRIFFSDIHFFGIQQNDDGSRRT TIVENVGSVEGLAYHRGWDILYWTYTSTTSTT RHTVDGTFGAFFRETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHFSINRAALSGANVL TLIERDIRTRGAAVGANKHLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTIEQGHVNCSGRGRILLQDDLTCRAVNSCR AQDSFECANGGCRINGLTCDGVPHCKDKSDE KPSYCNSRRCKETFRQCSNGCCVSNGM.WCN		l j					QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE
LPPIKL CDGADDCGGGSBEGLCDQCSLMN GGCSHNCSV APGGIVCSC JOMELGPDNHT CQIQSY CAKHLKCSQKCDQNKFSVKCSCYEG WYLEPDGESCRSLDPFKPFIIFSNRHERIDLH KGDYSYL/VPGLRNTIALDFHLSQSALYWTDV VEDKIYRGKLLDNGALTSEVVIQGLATPEG LAVDWLAGNIYWVESNLDQIEVAKLDGTLRT TLLAGDEHPRALALDPRDGILFWTDWDASLP RIEASMSGAGRRTVHRETGSGGWPNGLTVV DYLEKRILWIDARSDATYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGEVYWTDWTNTLA KANKWTGHNVTVVQRTNTQPFDLQVYHPSR QPMAPNPCEANGGQPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEKFLLYARQMER GYDLDAPYYNYDISTTVPDDNVTLUYDARE QRYYWSDVRTQAIKRAFBIGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQNNA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMAMMDGSNRTLLFSQQKGPVGL ADFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMBGUKWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGBCQLTPTSETTRSC MCTAGYSLRSQQQACEGVGSFLLYSVHEGR GIPLDPNDKSDALVPVSGTSLAVGIBFHAEND TITWVDMGLSTBRAKRDQTWREDVVTNGIG RVEGIAVDWLAGNIYWTDQGPDVIEVARLNG SFRYVVISGCLDKPRAITYDQGFPVEVARLNG GYPRIERSRLDGTGRVVLNVISISWPNGISV DYQDGKLYWCDARTDKERDLFTGENRVV LSSNNMDMFSVSVPEDPIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVPNRDR QKGTNYCAVANGGCQULKDIKVPNRDR QKGTNYCAVANGGCQULKDRKVPNRDR QKGTNYCAVANGGCQULKDRKVPNRDR QKGTNYCAVANGGCQULKDRKVPNRDR QKGTNYCAVANGGCQULKDRKVPNRDR RGSKDNATDSVPLRTGIGVQLKDRKVPNRDR QKGTNYCAVANGGCQULKTRSTTLIKSI HLSDERNLNAPVOPFEDEPHMKNVIALAFDY RAGTSPGTPNRIFFSDHFGNIQQNDDGSRRT IVENVGSVEGLAYHRGWDTLYWTSTTTTTI RHTYDQTRPGAFRETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLLEKDIRTNGLAPHRGLAFFALL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLLEKDIRTNGLAPHRGVGSMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTTHQWTOWCRARAVQSANKKLRYDDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTTHGRVNCSCRGRGRILQDDLTCRAVNSSCR AQDEFECANGECNSNLWCNN	1						GVTHVCDPSVKFGCKDSARCISKAWVCDGD
GGCSHNCSVAPGEGIVCSCPLGMELÖPDNHT CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG WYLEPGESCRSLDPFKPIFSNRHERRIDLH KGDYSVLVPGLRNTIALDFHLSQSALYWTDV VEDKLYKRKLLDNGALTSVLGTLRT LLAGDEIPRALALDFROEILFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRILWDARSDAIYSARYDOSGHMEVL RGHEFLSHPFAVILYGGEYWTDWRNTILA KANKWTGHNYTVVQRTNTQFFDLQVYHESR QPMAPNPCEANGGQOPCSHLCLDNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYYNYISTTVPDIDNVTVLDYDARE GRYYWSDVRTQAKRAFNGTOVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKQNNVA RLDGSFKNAVVQGLGPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGOKGPVGL ALDFFESKLYWISSGNHTRCNLDGSGLEVD AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVAHRMVYDESI QLDHKGTINFCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDFNDKSDALVPVSGOLCLPTSETTRSC MCTAGYSLRSQQACEGVGSFLLYSVHEGIR GIPLDFNDKSDALVPVSTSLAVVORGER GYRERSRLDGTGRAVVTNEDVVTNIGG RVEGIAVDWIAGNIYWTDQGFDVIEVARING SFRYVVISQGLDKPRAITVPEKGYLFWTEW GQYFRIERSRLDGTGRVVLVNYSISWNPGISV DYQDGKLYWCDARTNERBULTGENREVV LSSNNMMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGQULKDIKVPRIDR QKGTNVCAVANGGCQUCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSSTILAGSIK RGSKDNATDSVPLRTGIGQULKDIKVPRIDR QKGTNVCAVANGGCQUCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSSTILAGSIK RGSKDNATDSVPLRTGIGQULKDIKVPRIDR QKGTNVCAVANGGCQUCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSSTITAGSIK RGSKDNATDSVPLRTGIGQULKDIKVPRIDR QKGTNVCAVANGGCQUCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSSTITAGSIK RGSKDNATDSVPLRTGIGQULKDIKVPRIDR QKGTNVCAVANGGCQUCLYRGRGORACA CAHGMLAEDGASCREYAGYLLYSSTITAGSIK RGSKDNATDSVPLRTGIGQULKDIKVPRIDR QKGTNVCAVANGGCQUCLYRGRGORACA CAHGMLAEDGASCREYAGYLLYSSTITAGSIK RGSKDNATDSVPLRTGIGQULKDUKYPRIDR QKGTNVCAVANGGCQUCLYRGRGORACA CAHGMLAEDGASCREYAGYLLYSSTITAGSIK RGSKDNATGAVAVAGNAKKLRVOSMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCODLCL LTHGHVAVCSCRGGRILQDILTCRAVNSSCR AQDEFCANGECNPSLTCDGVPHCKDKSDE KPSYCNSRRCKKTRQCSNGKCNSNMLWCN GADDCGGOSDEIPCNKTAGGVGCFRCCRDGTC							NDCEDNSDEENCESLACRPPSHPCANNTSVC
CQIQSYCAKHILKCSQKCDQNKYSVKCSCYEG WYLEPOGESCRSLDPFISNRHERIDIH KGDYSVLVPGLRNTIALDFHLSQSALYWTDV VEDKIYRGKLLDNGALTSSEVVQYGLATPEG LAVDWIAGNIYWYSSNLDQIEVAKLDGTLRT TLLAGDEHPRAIALDPROELFWTDWDASLP RIEAASMSGGRRTVINESGGWPNGLTV DYLEKRILWDARSDAIYSARYDGSGIMEVL RGHELSHPRAVITYQETSGGWPNGLTV DYLEKRILWDARSDAIYSARYDGSGIMEVL RGHELSHPRAVITYQETSKFLLYARQMER QPMARNPCEANGGGPCSHLCLINYINTVSC ACPHLMKLHKDNITCYPFKKFLLYARQMER GYDLDAPYYNYIDSTVYDDDNYTVLYDARE GYPLYMSDVETQAKRAFNGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLQPHGLVYHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL ALDFESKL YWISSGNITINKCNLDGSGLEVID AMRSQLGKATALAMGGVSQLCLPTSETTRSC GTCSKADGSGSVULRNSTTLVMHIMKYYDESI QLDHKGTNPCSVNNGGCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GFLDPNDKSDALVPVSGTSLAVGIDFHAEND TITVWVDMGLSTISRAKRDQTWREIDVYTNGIG RVEGIAVDWIAGNITWTDQGFDVIEVARLNG SFRYVVISGGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISSWNGISV DYQDGKL WCDARTLKERNDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVHLTGIGVQLKDIKVPNRDR QKGTNVCAVANGGCQULLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFEDFEMIKNVIALAFDY RAGTSPGTPNRIFFSDHEGNIQQNDDGSRRTI VENVGSVEGLAYHRWHINNIALAFDY RAGTSPGTPNRIFFSDHERNIVALAFDY RAGTSPGTPNRIFFSDHERNIVQLBORGRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFEDFEMIKNVIALAFDY RAGTSPGTPNRIFFSDHERNIVALAFDY RAGTSPGTPRNIFFSDHERNIVALAFDY RAGTSPGTPNRIFFSDHERNIVALAFDY RAGTSPGTPRNIFFSDHERNIVALAFDY RAGTSPGTPRNIFFSDHERNIVALAFDY RAGTSPGTPRNIFF							LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
CQIQSYCAKHILKCSQKCDQNKYSVKCSCYEG WYLEPOGESCRSLDPFISNRHERIDIH KGDYSVLVPGLRNTIALDFHLSQSALYWTDV VEDKIYRGKLLDNGALTSSEVVQYGLATPEG LAVDWIAGNIYWYSSNLDQIEVAKLDGTLRT TLLAGDEHPRAIALDPROELFWTDWDASLP RIEAASMSGGRRTVINESGGWPNGLTV DYLEKRILWDARSDAIYSARYDGSGIMEVL RGHELSHPRAVITYQETSGGWPNGLTV DYLEKRILWDARSDAIYSARYDGSGIMEVL RGHELSHPRAVITYQETSKFLLYARQMER QPMARNPCEANGGGPCSHLCLINYINTVSC ACPHLMKLHKDNITCYPFKKFLLYARQMER GYDLDAPYYNYIDSTVYDDDNYTVLYDARE GYPLYMSDVETQAKRAFNGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLQPHGLVYHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL ALDFESKL YWISSGNITINKCNLDGSGLEVID AMRSQLGKATALAMGGVSQLCLPTSETTRSC GTCSKADGSGSVULRNSTTLVMHIMKYYDESI QLDHKGTNPCSVNNGGCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GFLDPNDKSDALVPVSGTSLAVGIDFHAEND TITVWVDMGLSTISRAKRDQTWREIDVYTNGIG RVEGIAVDWIAGNITWTDQGFDVIEVARLNG SFRYVVISGGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISSWNGISV DYQDGKL WCDARTLKERNDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVHLTGIGVQLKDIKVPNRDR QKGTNVCAVANGGCQULLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFEDFEMIKNVIALAFDY RAGTSPGTPNRIFFSDHEGNIQQNDDGSRRTI VENVGSVEGLAYHRWHINNIALAFDY RAGTSPGTPNRIFFSDHERNIVALAFDY RAGTSPGTPNRIFFSDHERNIVQLBORGRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFEDFEMIKNVIALAFDY RAGTSPGTPNRIFFSDHERNIVALAFDY RAGTSPGTPRNIFFSDHERNIVALAFDY RAGTSPGTPNRIFFSDHERNIVALAFDY RAGTSPGTPRNIFFSDHERNIVALAFDY RAGTSPGTPRNIFFSDHERNIVALAFDY RAGTSPGTPRNIFF			'			[GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT
KGDYSVLVPGLRYTTALDFHLSQSALYWTDV VEDKIYRGKLLDNGALTSEVVIQYGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TILLAGDIEHPRAIALDPRDGILFWTDWDASLP RIEAASMGAGRRTVHRETGSGGWPNGLTV DYLEKRILWIDARSDAIYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA KANKWTGHNYTVVQRTNTQPFDLQVYHPSR QPMAPPPCEANGQGGCSHLCLNYNRTVSC ACPHLMKLHKDNTTVEFKKFLLYARQMEIR GVDLDAPYYNYIISFTVPDDNYTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQNNA RLDGSFKNAVVQOLEQPHGLVYHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL ADPPESKLYWISSGRHTINCOLLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTILVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLFTSETTRSC MCTAGYSRSQQACGVGSFLLYSVHGGIR GIPLDPMDKSDALVPVSGTSLAVGIDFHAEND TYWVDMGLSTISRAKRDQTWREDVYTNGIG RVEGIAVDWIAGNIYWTDQGFDVEVARLNG SFRYYVISQGLDKFRATIYHPEKGYLFWTEW GQYPRIERSRLDGTERVVLNVNSISWPNGISV DYQDGKLYWCDARTDKERDLETGENREVV LSSNNIMDMFSVSFEDFTWSGTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKYPNRDR QKGTNVCAVANGGCQQLCLYRGRQGRACA CAHGMLAEDGASCRAYJLLYSERTLASI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDHFFNQQINDGSSRRILS IVENGSSVGGLAYBRGWDTLYWTSYTTSTIT RHITVDQTRRGAFERETVITMSGDHPRAFVL DECQNLMFWTNWRGJPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSRRYVLKSEPVHFGLAVYGEHIF WTDWVRRAVQRARHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGCCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTRQCSNGRLVNNSCR							CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
VEDKIYRGKLINGALTSFEVVIQYGLATPEG LAVDWIAGNIYWESNLOJEVAKLOGTI.RT TILAGDIEHPRAIALDPRDGILFWTDWDASLP RIEAASMSGAGRRYHRETGSGGWPNGLTV DYLEKRILWIDARSDAYRAYPOSGHMEVL RGHEFLSIPFAVTLYGGEVYWTDWRTINTLA KAMKWTGHNYTVVGNTQPPDLQVYHESR QPMAPNPCEANGGQGPCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIL GWYLDAYYNYNISTYPDIDNYTVLDYDARE QRVYWSDVRTQAKKAFINGTOVETVVSADL PNAHGLAVDWVSRNLFWTSVDTNKKQNVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMOSINRTLIFSGQKGPVGL AMPFESKLYWISSGRHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRSTLYMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQQACEGYSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDPHAEND TTVWVDMGLSTISRAKRQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDOGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDAKTDKIERDLETGENREVV LSSNNMDMFSVSYFEDFIYWSDSTHANGISK RGSKDNATDSVPLRTGIGVQLKDIKVYNRDR QKGTNYCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPTEDPEHMKNVALAFDY RAGTSPGTPNRIFSDHFGNIQQINDDGSRRIT IVENVGSVGGLAYIRGIGVLKDIKVYNRDR QKGTNYCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPTEDPEHMKNVALAFDY RAGTSPGTPNRIFSDHFGNIQQINDDGSRRIT IVENVGSVGGLAYIRGISWLAWTSTTSTIT RHTVDQTRPGAFERETVTIMSGDDHPRAFVL DECQNLMFWTNWNDQHPSIMFAALSGANVL TLIEKDIRTPNGLADDHRAEKLYFSDATLOKIE RCEYDGSIRYYULKSEEVHFGIAVYGEHIF WITDWVRRAVQRARKHVGSNMLURON JUTHQCHVNCSCRGGRILQDDLTCRAVYSGCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTRQCSNGRLQDDLTCRAVNSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTRQCSNGRCVSNMLWCN GADDCGOGSDEPECNKTACGVGETERCRUST					[WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH
LAVDWIAGNTYWESNLDQIEVAKLDGTLRT TILAGDIEHPRAIALDPROGILFWTDWDASLP RIEAASMSGAGRITVHRETGSGGWPNGLTV DYLEKRILWIDARSDAIYSARYDGSGHMEVL RGHEFLSHFFAVILYGGEVYWTDWRINTLA KANKWTGHNVTVVQRINTQPFDLQVYHPSR QPMAPNPCEANGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKPLLYARQMEIR GVDLDAPYYNYIISTVPDIDNYTVLDYDARE GRYVSSDWRTQAKAFNINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQNNA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMAMMDGSNRTLLFSGQKGPVGL ADPFESKLYWISSGNHTINKCNLDGSGLEVID AMRSQLGKATALAMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGDDFHAEND TTYWVDMGLSTSRAKRDQTWREDVVTNCIG RVEGIAVDWIAGNIYWTDQGFDVEVARLNG SFRYVVISQLDKPRAITVHPEKGYLFWTEW GQYPRERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKERDLETGENREVV LSSNMDMFSVSYEDFTYWSDRTHANGSK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPHMKNVIALAFDY RAGTSPGTPNRIFFSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRFGAFERETVITNSGDDHFRAFVL DECQNLMFWTNWNFQHFFSIMKALGSANVL LIJEKDRTPRGLAFDHFRGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRFGAFERETVITNSGDDHFRAFVL DECQNLMFWTNWNFQHFFSIMKALGSANVL LIJEKDRTPRGLAFDHFRGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRFGAFERETVITNSGDDHFRAFVL DECQNLMFWTNWNFQHFFSIMKALGSANVL LIJEKDRTPRGLAFDHFRGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRFGAFERETVITNSGDDHFRAFVL DECQNLMFWTNWNFQHFFSIMRALGSANVL LIJEKDRTPRGLAFDHFRGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRFGAFERETVITNSGDDHFRAFVL DECQNLMFWTNWNFQHFFSIMFALGSANVL LIJEKDRTPRGLAFDHFRGNIQQDHCTGAVNSCG AQDEFCANGECNFSLTCDGVPHCKDKSDE KFSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGGSSEDFCNGTGCTCSNMLWCN	1	1					KGDYSVLVPGLRNTIALDFHLSQSALYWTDV
TILLAGDIEHPRALALDPROELIFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRIL WIDAASDAIYSARYDGSGHMEVL. RGHEFI.SHPFAVTLYGGEVYWTDWRTNTLA KANKWTGHNVTVVQRTNTQFPDLQVYHPSR QPMAPNPCEANGGQGPCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKFLLVARQMEIR GVDLDAPYYNYIISTVPDIDNVTVLDYDARE QRVYWSDWRTQALKRAFNGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLQPHGLVVHPLRGKLY WTDGDNISMAMMDGSNRTILFSGQKGPVGL ADFPESKLYWISSGNHTILRSGQKGPVGL ADFPESKLYWISSGNHTILRSGQKGPVGL AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNGDCSQLCLPTSETTRSC MCTAGYSLRSQQQACEGVGSFLLVSVHEGIR GPLDPNDKSDALVPVSGTSLAVGIDFHAEND TITVWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQFDVEVARLNG SFRYVVISQGLDKPRATTVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DVQDGKLYWCDARTDKIERDLETGENREVV LISSNMDMFSSVYEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQJCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVALAFDY RAGTSGTPRNIFFSDHFGINQQINDDGSRRIT VENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHITVDQTTRGAFFERETVITMSGDHFRAFVL DECQNLMFWTNWNEQHFSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVLKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVOSNMKLLEVDIPQ QPMGILAVANDTNSCELSFCRINGGCQDLCL LTHQGHVNCSCRGGLQDDLTCRAVNSSCR AQDEFSCANGECNFNLTCDGVPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNM.WCN GADDCGDGSDEIPCNKTACGVGEFRCLWCN	i		-				VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG
RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRILWIDARSDAJYSARYDGSGHMEVL RGHEFLSHFAVTLYGGEVYWTDWRTNTLA KANKWTGHNYTVVQRINTOPFDLQVYHPSR QPMAPNPCEANGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKPLLYARQMEIR GVDLDAPYYNYISTTVPDIDNYTVLDYDARE QRYVWSDWRTQAURAFNOGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQHEIQVHPLRGKLY WTDGDNISMANNDGSRRTLLFSGOKGPVGL ADPFESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGIDHHAEND TIYWVDMGLSTISRAKRDQTWREDVYTNGIG RVEGIAVDWIAGNYTDQGFDVEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNYSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSVSVFEDFYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVLKDIKVFINDR QKGTNVCAVANGGCQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVALAFDY RAGTSSGTPNRIPFSDHFGINQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVQOTRPGAFRETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLJEKDRTPRIGLADHRAEKLYFSDATLDKIE RCEYDGSRRYVILKSEPVHPFGLAVYGEHF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGCDNSTLTCGGVGEFRCRUGTC							LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT
DYLEKRIL WIDARSDAIYSARYDGSGHMEV.I RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA KANK WTGHNYTVVQRTNTQPFDLQVYHPSR QPMAPNPCEANGGQGPCSHLCLINYNRTVSC ACPHILMCHIKDNTTCYEFKKFLLYARQMEIR GVVLDAPYYNYIISFTVPDIDNVTVLDYDARE QRYVWSDVRTQAIKRAFINGTGVETVYSADL PNAHGLAVDWYSRNI-FWTSYDTNKKQNNA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQACEGVGSFLLYSVHEGIR GPLDPDNDKSDALVPVSGTSLAVGDPHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVUEVARLNG SFRYVVISQQLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKHERDLETGENREVV LSSNNDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAV ANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDHIFGNIQQUNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLJEKDIRTPNGLADHRAEKLYFSDATLDKIE RCEYDGSHRYVNSEPVHPRGLAVYGEHF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCGGGRILQDDLTCRAVNSSCR AQDEFGCANGECNFSLTCDCVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN	1						TLLAGDIEHPRAIALDPRDGILFWTDWDASLP
RGHEFLSHPFAVIL YGGEVYWTDWRTNTLA KANKWTGHNVTVVQRTNTQFFDLQVYHPSR QPMAPNPCEANGQGPCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYFFKFLLYARQMEIR GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE QRYYWSDVRTQAKRAFINGTGYETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQNNA RLDGSFKNAVVGOLEQPHGLVYHPLRGKLY WTDGDNISMANMDGSNRTLLFSQKGPVGL AIDFFESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAMGDKL WADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSQQACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRQTWREDVYINGIG RVEGIAVDWIAGNTYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATIDSVJFEDFTYWSDRTHANGSIK RGSKDNATIDSVJPEDFEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWOTLLYWTSYTTSTIT RHTVDQTRPGAFRETVITIMSGDDHFRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLJEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYTDGSHTYVILKSEPVHPFGLAYYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTINSCELSPCRNINGSCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECNFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLLWCN GADDCGDGSSBEPCNKTACGVGEFRCRUGTC	1						RIEAASMSGAGRRTVHRETGSGGWPNGLTV
KANK WTGHNYTVVQRTNTOPFDLQVYHPSR QPMAPNPCEANGGQGCSHLCLNYNRTVSC ACPHLMKLHKDNTTCYEFKKPLLYARQMEIR GVDLDAPYYNYIISFTVPDIDNYTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWYSRNLFWTSYDTNKKQNNA RLDGSFKNAVVQGLEOPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISGRHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGGSGVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVYEGIR GPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVYTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSVSYFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCRYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTTPRRIFFSDIHFGNQQNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVTIMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TILIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVLKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHYGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRNNGGCQDLCL LTHQGHVNCSCGGRILQDDLTCRAVNSSCR AQDEFGCANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLUCN GADDCGDGSBEPCNKTACGVGEFRCRDGTC							DYLEKRILWIDARSDAIYSARYDGSGHMEVL
QPMAPNPCEANGGQÖPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYYNTISTIVPDIDNVTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLWHHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVGGTSLAVGIDFHAEND TIVWDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSIADTERVVLNVSISWYPMGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYJLLYSERTILKSI HLSDERNLNAVQFFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDHFGNIQQDNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRRGAFERETVITMSGDDHFRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSIRPYLKSEPVHFFGGAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGGVGEFRCRDGTC							RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA
ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYYNYISTIVPDIDNVTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVYSADL PNAHGLAVDWVSRNLFWTSYDTINKKQDNVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMAMMDGSNRTLLFSGGKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPIDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWLAGNIYWTDQGFDVIEVARLNG SFRYVIVSQGLDFRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKERDLETGENREVV LSSNNMDMFSVSVPEDFTYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLLYFSDA TLDKIE RCEYDGSIRRYVILKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANGROCQLCL LTTIQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGYGEFRCRDGTC							KANKWTGHNVTVVQRTNTQPFDLQVYHPSR
GVDLDAPYNYIISTVPDIDNVTVLDYDARE QRVYWSDVRTQKTAINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMGSNRTLLFSGQKGPVGL ADFPESKLYWISSGNITTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGDFHAEND TIYWVDMGLSTISRAKRQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATDSVPLATGIGGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLADHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCEL SPCRNNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN		}					QPMAPNPCEANGGQGPCSHLCLINYNRTVSC
QRVYWSDVRTQALKRAFINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL ADFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGYSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKRRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSYFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGQQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYYLLYSERTILKSI HLSDERNLNAPVQPEDPEHMKNVIALAFDY RAGTSPGTTPNIFFSDHIFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDA TLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGGDSDEIPCNKTACGVGEFRCRDGTC						i	ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR
PNAHGLA VDWYSRNLFWTSYDTNKKQINVA RLDGSFKNAVVGGLEQPHGLVYHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQFFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDHFGNIQDNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRRGAFERETVITMSGDDHPRAFVL DECQNLMFWTNNWEQHPSIMRAALSGANVL TILIEKDIRTPNGLADHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	i						GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
RLDGSFKNAVVQGLEQPHGLVVHPLRĞKLY WTDGDNISMANMDGSNSTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGLAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGORACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1						QRVYWSDVRTQAIKRAFINGTGVETVVSADL
WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFFESKLYWISSGNHTINRCNLDGSCLEVID AMRSQLGKATALAMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLFTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNTYWTDQGFDVIEVARLNG SFRYVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRITI TVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRFGAFERTVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSINRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHFFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC						·	PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
ADFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKL WWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKYYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGLAVDWLAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQUCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTYDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECNPSLTCDGVPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEPCNKTACGVGEFRCRDGTC							
AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKAADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVLLKSEPVHPFGLAVJGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRNNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTAGGVGFERCRDGTC)						WTDGDNISMANMDGSNRTLLFSGQKGPVGL
GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHFRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHYGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC							AIDFPESKLYWISSGNHTINRCNLDGSGLEVID
QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVLALAFDY RAGTSPGTPNRIFFSDHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVYHCKUSDE KPSYCNSRRCKKTFRQCSNGRCVSNMI.WCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1						AMRSQLGKATALAIMGDKLWWADQVSEKM
MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1				ł		GTCSKADGSGSVVLRNSTTLVMHMKVYDESI
GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND TTYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRQCRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHITVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRNNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVYPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1						QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC
TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWINWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					l		
RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			1		-		
SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	'					1	
GQYPRIERSRLDGTERVVLVNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC				-	l		
DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC				l	ł		
LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHFRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					ļ		
RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					1		
QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			}		l	i	
CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			1				
HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC					1		
RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			ł	1	ŀ	ł	
IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC							
RHTVDQTRPGAFERETVITMSGDDHPRAFVL DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			ļ	ļ		}	
DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGILAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			- 1	Ì			
TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			ļ				
RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			- 1			Ì	• • • • • • • • • • • • • • • • • • • •
WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			ſ			ľ	
QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC							
LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	1						
AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC			l	- 1	1		
KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC	.	ļ	1				-
GADDCGDGSDEIPCNKTACGVGEFRCRDGTC		İ	· ·	i	1	,	
	1	1	ľ	ĺ	1	,	
		ļ	Į		1	·	
			1	}			IGNSSRCNQFVDCEDASDEMNCSATDCSSYF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
						nucleotide insertion
\vdash				sequence		L
						RLGVKGVLFQPCERTSLCYAPSWVCDGAND
1						CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP
						MSWTCDKEDDCEHGEDETHCNKFCSEAQFE
1					1	CQNHRCISKQWLCDGSDDCGDGSDEAAHCE
						GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA
						DGADESIAAGCLYNSTCDDREFMCQNRQCIP
						KHFVCDHDRDCADGSDESPECEYPTCGPSEF
1			1			RCANGRCLSSRQWECDGENDCHDQSDEAPK
1						NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
1						GQDDCGDSSDERGCHINECLSRKLSGCSQDC
						EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
						TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP
						HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY
1						TLLKQGLNNAVALDFDYREQMIYWTDVTTQ
1						GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV
1						DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL
						VSSGLREPRALVVDVQNGYLYWTDWGDHSL
						IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE
						RIYWADAREDYIEFASLDGSNRHVVLSQDIPH
						IFALTLFEDYVYWTDWETKSINRAHKTTGTN
1						KTLLISTLHRPMDLHVFHALRQPDVPNHPCK
1						VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRICUSNCTASOFUCKNDKCIPFWWKCDIF
						GRTCVSNCTASQFVCKNDKCIPFWWKCDTE
						DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN
		١.,		•		PAFICDGDNDCQDNSDEANCDIHVCLPSQFK
						CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE
						VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
!						GSDEPANCTQMTCGVDEFRCKDSGRCIPARW
						KCDGEDDCGDGSDEPKEECDERTCEPYQFRC
1						KNNRCVPGRWQCDYDNDCGDNSDEESCTPR
]		PCSESEFSCANGRCIAGRWKCDGDHDCADGS
						DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA
1						DADCMDGSDEEACGTGVRTCPLDEFQCNNT
				·		LCKPLAWKCDGEDDCGDNSDENPEECARFV
						CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD
						GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
1						SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT
						NASICGDEARCVRTEKAAYCACRSGFHTVPG
						QPGCQDINECLRFGTCSQLCNNTKGGHLCSC
						ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS
1						LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA
						GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR
]						RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY
1						WTDSGRDVIEVAQMKGENRKTLISGMIDEPH
1						AIVVDPLRGTMYWSDWGNHPKIETAAMDGT
1						LRETLYODNIOWPTGLAVDYHNERLYWADA
]						KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
			[ĺ	FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG
					į	_ · · · · · · · · · · · · · · · · · · ·
				l		LSHASDVVLYHQHKQPEVTNPCDRKKCEWL
						CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP
					ļ	PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC
(l	-	1	ľ	ì	QPRYTGDKCELDQCWEHCRNGGTCAASPSG
					ļ	MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT
						VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE
				1		NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
					ŀ	CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
		ļ				CLTCVGHCSNGGSCTMNSKMMPECQCPPHM
						TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL
) . I						VAGVVFWYKRRVQGAKGFQHQRMTNGAM
				·		NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	}	Ì	Į.	peptide	1	/=possible nucleotide deletion, \=possible
L		<u></u>		sequence		nucleotide insertion
			1			DPDKPTNFTNPVYATLYMGGHGSRHSLASTD
	ļ		l	L		EKRELLGRGPEDEIGDPLA
374	1724	A	3187	191	1815	CLELASAGKIPEESKALSLLAPAPTMTSLMPG
	·					AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI
1				ļ		ATLGEIPQPPLMGNVDPSKIDEIRRTVYVGNL
1	į	ŀ			}	NSQTTTADQLLEFFKQVGEVKFVRMAGDET
			ļ			QPTRFAFVEFADQNSVPRALAFNGVMFGDRP
1	i					LKINHSNNAIVKPPEMTPQAAAKELEEVMKR
			İ		l	VREAQSFISAAIEPGWLHSTSLCNDFLGCF*RR
·{			[ĺ	RMYRE*APCTICGTFHLCLIINWDL*LF*AYTA
						K*FFPPRVWKEQ*KKRR\RSRSHTRSKSRSSSK
		1]		}	SHSRRKRSQSKHRSRSHNRSRSRQKDRRRSK
1					1	SPHKKRSKSRERRKSRSRSHSRDKRKDTREKI
						KEKERVKEKDREKEREREKEREKERGKN
1						KDRDKEREKDREKDKEKDREREREKEHEKD
						RDKEKEKEQDKEKEREKDRSKEIDEKRKKDK
1						KSRTPPRSYNASRRSRSSSRERRRRRSRSSSRS
					•	PRTSKTIKRKSSRSPSPRSRNKKDKKREKERD
						HISERRERERSTSMRKSSNDRDGKEKLEKNST
375	1725	Ā	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG
313	1725	^	3172	413	101	PFPNLKKSLRGTHFSSVKK\TTLTWLNSODP
						WF/FFYP*SPDLQIPSSFRNGLNDWYHHSQKC
						PDLDGAYVKK
376	1726	Α	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR
370	1720	^	3177	931	1 410	HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS
1						ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI
		!				K
377	1727	A	3201	274	1285	KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM
"			520.		1205	GYKLQDLTDVQIMARLQEESLRQDYASTSAS
]			VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE
·						EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC
1						RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP
		İ				NRTNGDK/PPKKYA*PSPDAKYNCH**QH\SSP
						VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL
		۱ ;				QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS
						NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS
						ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL
					·	SSLSTLRDGNWRDGCY
378	1728	Α	3202	112	1789	VPGVTESRPSVLRGDHLFALLSSETHQEDPIT
						YKGFVHKV\ELDRVKLSFSMSLLSRFVGWG*
						PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW
						PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP
						EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT
i i				' I		LVEAIKQVVKHLPKAHILACAPSNSGADLLC
						QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN
	į					WDAKKGEYVFPAKKKLQEYRVLITTLITAGR
						LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG
						LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL
						TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ
				·		FITKLLRNYRSHPTILDIPNQLYYEGELQACA
1 1			İ	ł	l	DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD
	1				l	EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK
						GKARLSPRSVGVISPYRKQVEKIRYCITKLDR
j	1		ĺ		1	ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP
1			[l		ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD
						SNLRVWDGIRKPACLTNTSCHS
379	1729	Α	3206	432	130	PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK
<u></u>				i		*LSTREAXDSXPGRQIAXXRQGGKVETTTAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted begiming nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion XKOSNNKGTRASSYXEPDAXEOWKFPHKKL
300	1720	A	2207	107	507	QLPGXTHE GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS
380	1730		3207	187	307	PEQRDHPAPHKQIEAGQGLPGPQAWGG*KGP AXLLPGPGGGPGPVASLEARAQASSGVTPNG GGRTYPYPTFSSGE
381	1731	A	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/ EMQLITSLGLQEFDIARNVLELIYAQTLVWIGI FFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPS KAWRASQMMTFFIFLLFFPSFTGVLCTLAITI WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP GYLWVWIYRNLIGSVHFFFILTLIVLITTYLY WQITEGRKIMIRLLHEQIINEGKDKMFLIEKLI KLQDMEKKANPSSLVLERREVEQQGFLHLGE HDGSLDLRSRRSVOEGNPRA
382	1732	A	3238	256	38	LLMIKVSSTCFSCHLHHHHHHHHHRHHQGHNS LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF LFSNACL
383	1733	A	3241	1542	343 678	KGAPSFVRLYQYPNFAGPHAALANKSFFKAD KVTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLPKNGPIYDVVWNS SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP
						RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI
385	1735	A	3243	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PPPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \mathref{mathreo}possible nucleotide insertion MTCADPGEIANGHRTASDAGFPVGSHVQYRC
						LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC ALKYEPCLNPGVPENGYQTLYKHHYQAGESL RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF SNPLYEAGDTREYEVSI
386	1736	A	3250	5725	3984	GTSTVTMATKKHFSIILNLLGMLLKKDNQDT RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTI.VEDVNICLQACSSLHALSSS LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP SNIFHPQDFSD/VISFILYGNSHRTGKDNWLE RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI MMVVEALCELHCPEAIQGIAVWSSSIVGKHL LWINSVAQQAEGRFEKASVEYQEHLCAMTG VDCCISSFDKSVLTLASAGCKSASLKHCLNGE SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF NYIKSLSSFESGKFVECTEQLELLPGENINLLA GGSKEKIDMKKLLRNM
387	1737	A	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF LKYVHNL*AENYKTLMK*INEDLNKQRDVPY S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMQ
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG FELLDSSDLPASASKSAGITCMSHHARTLSLK *WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI LTRLETQMINADYQNKLTLDYLLTTDREVYE PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVQV*HGFDPEAMFR
390	1740	A	3270	2	372	GRCHDQNKGKS\DGPDAQAEACGGESTYQEL LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1	187	FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ HVETFFQ\EELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY PVLKEDAPRQHVETFFQVGINQKSRGHEVRR KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVLPRQANFCIF\M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFALKGCLPRQKEGGTLNLI
393	1743	A	3283	385	3	RNRSVVPEFVLLGLSAGPQTQTLLFVLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM A*VFFVFATGGTESSLLAVMAYDRYVAIRTR G
394	1744	A	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCQKGERGKKGRE

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide	1	/=possible nucleotide deletion, \=possible
		1		sequence	1	nucleotide insertion
						RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN
						KQQQ
395	1745	Α	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
		1			Ì	WICLSMVILTHSLKTFHRNWDWESEYTLFMS
	1					ALKVNKNNAKLWNNVGHALENEKNFERAL
L	<u> </u>					KYFLQATHVQPDDIGAHMNVGR
396	1746	Α	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL
<u> </u>						IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	Α	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL
1						WSEACAFL*AAAPQGPASPCCGLPSGFPRVW
			1			AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS
			1			GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA
<u></u> _						GLT
398	1748	A	3300	1912	2768	KQRRWQNIQRKGPKRYTVIAGNSQSHQPMIFS
	l	!				MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY
						STSFLTDSYLKYIGWTLHDKHREVRVKCVKA
		ł	ľ			LKGLYGNRDLTARLELFTGRFKDWMVSMIV
1		ł	1	-		DREYSVAVEAVRLLILILKNMEGVLMDVDCE
İ	ľ					SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI
						RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH
1		1				SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR
	1	1	i • 1			VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL
}						VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHK\FLFLLLPSLLMGYSE
		}				SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI
						CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI
						PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS
						GRAVALLHLIASGLTSIQTNTASSKPPIWGYL
			1			STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ
						SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV
	ł					LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN
1	•					SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS
						N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL
		1				TGAALAGSYPIWENENTLSWLPTFTYNFCLST
ļ						PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI
1	1		1			NILPPNQTILISVEASISSSPIRNKWALHLITLLT
						GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE
						DMHTSITSLQRQLDFLVGVILQNWRVLDLLT
						TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH
]			DRAAEL*HQVADSWWQGSSLLRWIPWVAPF
		1				LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ
].				ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR
						P
400	1750	A	3303	2	453	THWRHSSGVPGSTTARRRRELEIATSDNQE
1						YYNRLCQEVTNRERNDQKMLADLDDLNRTK
						KYLEERLIELLRDKDALWQKSDALEFQQKLS
						AEERWLGDTEANHCLDCKREFSWMVRRHHC
						RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ
-		•			_	CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP
						EKIVLRALKDSRAGMPEQDKDPGVQENPDD
[]	QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP
•						GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR
						NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG
						LSLSTQILG*QKPSKYIPSLCKR
402	1752	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA
		- 1				PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF
						VVRTMCAVLGLVARQEDSGLRDHSVRVLISN
				•		HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ FFFFIFLNIFILAFSSPGSQPLLNSPPSFVCWSR GFMEMNGRGELVESLKRFCASTRLPPTPLLLF PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ VQRTLVSVTVSDASWVSELLWSLFVPFTVY QVRWLRPVHRQLGEANEEFALRVQQLVAKE LG\(QTGTRLTPA\)DKAEHMKRQRHPRLRPQS AQSSFPPSPWVLSS\('SDVQTGQTLGFREFKESF CPHVAIGVFIPERPWPKTGCCKTLTIHLILL*G GPVSFSCPEDIHPRGT*VPTQQASGLPSFPSYG PARGGVL*HPSAQQPLTFAKSS\('WARAGRAL
L						QERKQ\ALYEYARRFTERRAPGGLD
403	1753	A	3307	44	447	DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS GGASAGLASSPECACGRSHFTCAVSALGECT CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD CPPRECEED
404	1754	A .	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG QDHVQNEEIYARVLDKFGSNFLSRDNADLGT AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS LLKGDLKGVKGDLKKPFDKAWKDYETKFAK IEKEKREREWR
405	1755	A	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG HPASKENDQMVDTIKNTTKVPIIWTYGDMVE PRPQMIRPAVGAKHKELWKILMALKKIKIWE GKYTKPSQYNPNYMLELAHNDSVW
406	1756	A	3324	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV MCVLLWALSLLQSILEWMFCSFLFSDVDSDN WCQILDFLTAVWLIFLILVLCGFTLVLLVRIIC GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F LLYWIEKDLDDL
407	1757	A	3328	213	1841	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC PHISFTGNICVYCPGGPDSDFEYSTQSYTGYEP TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD KVEFIVMGGTFMALPEEYRDYFIRNLHDALS GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC MKRHLSDMLTYGCTRLEIGVQSVYEDVARD TNRGHTVKAVCESFHLAKDSGFKVVAHMMP DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP TLVIRGTGLYELWKSGRYKSYSPSDLVELVA RILALVPPWTRVYRVQRDIPMPLVSSGVEHG NLRELALARMKDLGIQCRDVRTREVGIQEIH HKVRPYQVELVRRDYVANGGWETFLSYEDP DQDILIGLIRLRKCSEETFRFELGGGVSIVREL HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL QGPYMVKMLK
408	1758	A	3335	3	467	AIASPRAAGIRHELTSTMAAGKNKRLTKGGK KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG LKGRVFEESLADLQND\TDGYLLRVI*VAFTT ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH DDFARKVKMLKKPKFELRKLMELHGEGSS
409	1759	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK KGIRARILETLGMLLLLALLILGIVWVASALID NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- cotide scq- uence	SEQ ID NO: of peptide seq- uence	Met. hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERKKASAWERNLVYPA VMVLLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK
410	1760	A	3339	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLÜP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342		2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVVKAGESYTHIQYLFQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPRDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVPSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	RRYEKFOK VLIALSHKLEFAVRSSEL IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VMPLVRMPWKRAVVLLMLWFIGQAMWLAP AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII SHYKEEPLTERIKYD PIPVRWNSLEGRLLRGYEQHANDGKDYISRN *DLRSWTAADMAAOITKRKWEAEEFAEOIKA
		· !				YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS GHFLPTDRVSYSEAASSDHAQGSDVSLTACK V
414	1764	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL GVHMVDKDTERDIEMKRQLRRLRELHLYST WKKYQEAMKTSLGVPQRERDEGSLGKPLCP PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP EASNQSLLTVAHADAGTQTNGDLEDLEEHGP GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEEEDPEEVKTSLGVPQRGDLEDLE EHVPGQTVSEEATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEEATGVHMMQVDPATLA KQLEDSTITGSHQQMSASPSSAPAEEATEKTK VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD IFNIF
415	1765	A	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH
417	1767	A	3382	2	2061	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLKLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAGAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	Α .	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NKVHADLVISKPVSKSPERLRKDIEVLSEDTD YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR RYCNTEECLKTGSPGKKEEKAKNKESLCMEN SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK SLRTTGFYSGFSEVAEKRIKLLNNSDERLQNS RAKDRKDVWSSIQGQWPKKTLKELFSDSDTE AAASPPHPAPEEGVAEESLQTVAEEESCSPSV ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE LQDLQSERE*LASRF*CQCELKQ**SARTRTS*KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK
419	1769	A	3399	206	463	QQKEGK QRECLSIHIGQAGIQIGDACWELYCLEHGIQP
						NGVVLDTQQDQLENAKMEHTNASFDTFFCE TRAGKHVPRALFVDLEPTVIDGIR
420	1770	A	3408	1010	685	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV VMGFHHVGQAGLELLTSGDLPALASQSARIT GVNHCAQPRGHFH
421	1771	A	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL LYLVSPLENEPKEMLTLSEYHERVRSQGQQL QQLQAELDKLHKEVSTVRAANSERVAKLVF QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRNTAYFWNRFSFWNYARPPTVILEPHVFP GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT SEGAEGSAQGPH
422	1772	A	3412	2	421	EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN YLVSCSMRSPESGKGEPGTARDYTPMGRPPP PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP VDTAGAPASPGPDVCE
423	1773	A	3420	91	706	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/ RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS KTHLPGFVEQAEALKAKGVQVVACLSVNDA FVTGEWGRAHKAEGKVRLLADPTGAFGKET DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA LNVEPDGTGLTCSLAPNIISQL
424	1774		3421	4	7688	RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL VTVQVLDINDNAPIFVSTPFQATVLESVPLGY LVLHVQAIDADAGDNARLEYRLAGVGHDFP FTINNGTGWISVAAELDREEVDFYSFGVEAR DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE RQYVLAVTASDGTRQDTAQIVVNVTDANTH RPVFQSSHYTVNVNEDRPAGTTVVLISATDE DTGENARITYFMEDSIPQFRIDADTGAVTTQA ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI LVNDVNDNAPQFLRDSYQGSVYEDVPPFTSV LQISATDRDSGLNGRVFYTFQGGDDGDGDFI

NO: of No: of operation of the control of the contr	CONC.	1 000					
umce entide seq- umce unce cut (USN) 094905 094905 094905 094905 094906	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence USSN Ocation Sequence		i	nod				, •
uence 1914 0.0 miss amino acid cristion of peptide residue of peptide sequence propried sequence propried sequence 0.0 miss amino acid cristion of peptide sequence 0.0 miss amino acid cristion of peptide sequence 0.0 miss amino acid cristion of peptide sequence 0.0 miss amino acid cristion 0.0 miss amino acid cristion 0.0 miss amino acid cristion 0.0 miss amino acid cristion 0.0 miss amino acid cristion 0.0 miss amino 0.0 miss amino acid cristion 0.0 miss amino 0.0		1					
umoc 1914 maino aedi meido aedi residue of peptide peptide peptide sequence T-Thronoine, V-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryprosine, Y-Valina, W-Tryptophan, Y-Tryp	1	1 .	ĺ				
mino seid reduce of peptide sequence pin		uznce				1	,
residue of popidid sequence Variance Va	uence	l	l	714			
peptide sequence Possible nucleotide deletion, impossible nucleotide insertion VESTSGIVETT.RR.DRENNAQVVLRAYADDK GMPARTPMEVITVLUDNIDNIPPYEQDETD VFEENSPIGLAVARVTATIPDEGTTAQMY QIVEGNIPPYEQUETEGELTALVID DEPENPE VFVEENSPIGLAVARVTATIPDEGTTAQMY QIVEGNIPPYEQUETEGELTALVID DEPENPE VILVIQATSARP VSRATVHIVRLLDRENDIPPY LINESTLENNIVTINESSSFPOGLAREVALIDP DISDSLTYSFERGNELS.VILLNASTGELKLSR ALDINNEPLEAMS VLYDOVHSYTAQCALRV TITTDEMLTHSTITLEDMSPERTLSPLICLFIQ AVAATALTPEDMV PRVQRUTDARGGHILN VSLSVQPPGPGGGPFFLSEDQUET, VLNRS LITAISAQVPTPUDMICLEPCENYMECVSV TITTDEMLTHSTITLEDMSPERTLSPLICLFIQ AVAATALTPEDMV PRVQRUTDARGGHILN VSLSVQPTQPGGGPFFLSEDQUET, VLNRS LITAISAQVSTPODMICLEPCENYMECVSV CLREDOTTORHERS SASTRASSON, PREPRIPPYOGLERCYPOF TIDYCETEVDLCYSR.PGHPRGCRSREGGYT CLREDOTTORHERS SASTRASSON, PREPRIPPYOGLERCYPOF TIDYCETEVDLCYSR.PGHPRGCRSREGGYT CLREDOTTORHERS SASTRASSON, PREPRIPPYOGLERCYPOF TIDYCETEVDLCYSR.PGHPRGCRSREGGYT CLREDOTTORHERS SASTRASSON, PREPRIPPYOGLERCYPOR TOYCETEVDLCYSR.PGHPRGCRSREGGYTTRSSASTRASSON, PREPRIPPYOGLERCYPOR TOYCETEVDLCYSR.PGHPRGCRSREGGYTTRSSASTRASSON, PREPRIPPYOR TOYCETEVDLCYSR.PGHPRGCRSREGGYTTRSASTRASSON, PREPRIPPYOR TOYCETEVDLCYSR.PGHPRGCRSREGGYTTRSASTRASSON, PREPRIPPYOR TOYCETEVDLCYSR.PGHPRGCRSREGGYTTRSASTRASSON, PREPRIPPYOR TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TOYCETEVDLCYSR.PGHPRGCRSREGGYTH TO	1	!	1]			
Intelestic insertion		Ì	1			sednence	
VESTSGIVETLIRRIDRENYAQYVILRAYADIS GMPARTPMEVITU/LDVINDIPYEQDEED VFVEENSPIGLAVARVTATIPPEGTEAQIMY QIVEGINIPYPVQLIDIFSGELTALVILDJYEDREE YVLVIQATSAPLVSRATVIKYRLIDRIDRIPYP LOMFELLENNYVINESSSFPGAGEISLYLLINASTGELKLISR ALDNIRPLEAIMSVLVSDGWHSVTAGCALRV TITTDEMLTHSTITRLEDMSPERTLSPLLGLEIQ AVAATLATPPINLYVNNQDDTHAYACCALRV TITTDEMLTHSTITRLEDMSPERTLSPLLGLEIQ AVAATLATPPINLYVNNQDDTHAYACCYSV LERDSSAPTASSSVLERPETPVGGLRCRCOPGE TGDYGETEVDLCYSRASGRCTPGVGNGGT CONLLVGGRCDGSGFENSEDLQERLVILNES LITAISAGRVLPEDDIOLICLEPCENSYMCGVSV LERDSSAPTASSSVLERPETPVGGLRCRCOPGE TGDYGETEVDLCYSRASGRCTPGVGNGGT CYNLLVGGRCDGTSGASSSVLTRETETPVGGLRCRCOPGE TGDYGETEVDLCYSRASGRCTPGVGNGGT CYNLLVGGRCDGTSGASSSVLTRETETPVGGLRCRCOPGE TGDYGETEVDLCYSRASGRCTPGVGNGGT CYNLLVGGRCDGTSGASSSVLTRATETPVGGLRCRCOPGE TGDYGETEVDLCYSRAGGRCTPGVGNGGT CYNLLVGGRCDGTSGASSSVLTRATETPUGGLRCRCOPGE TGDYGETEVDLCYSRAGGRCTPGVGNGGT CYNLLVGGRCDGTSGASGAGGATTAGGTG CYNLDAGGRCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1		ĺ	1	• •		
GMPPARTMENTYTULDVINDIPPVERQDEID VVEENSPIGLAVARTATDPEGTHAQMY QIVEGNIPEVPQLDIPSGELTALVDLDYEDREE YLVLQATSAPLVSRATVINLALDRINDIPPV LGNFELLFINTYTTRSSSFEGAGREVPAHDP DISDSLTYSFERGNEIS.VLINAASTGELKLSR ALDNRRPLEAMSVLVSDGVHSVTAQCALRV ITITDEM THSTITLALDMSPERLSPLQLIFLQ AVAATLATPDHVVVFNVQBOTDAPGGHLIN VSLSVGOPPOPGGGFPPSEDLOGELVLNRS LLTAISAQRVLPPDDNICLEPCENYMECVSV LEPDSSAPFLASSVLAFPIPPVGGLRCCPPGE TGDYCETEVPLCYSRPOPHGRCRSRGGGT CONLLVGGFKCDCFSGPFEXPCQVGTTGCYPGF TGDYCETEVPLCYSRPOPHGRCRSRGGGT CONLLVGGFKCDCFSGPFEXPCQVGTTGSFP ARISTITEGERQFHFTLALSFATKERDGLLL YNGRRINGKHDFVALEVIGQVQLTFSAGEST TTVSPFVPGGVSDQWHTVQLKYYNKPLLG GGLPGPSEQKLAVVTVOCCTGVALEFGS VLGNTSCAAQGTGGGSKKSLDLTGFLLLGG VDLPESPFVBMRGFVGCKSRLQWAVTSGCTGTGVALFGS VLGNTSCAAQGTGGGSKKSLDLTGFLLLGG VDLPESPFVBMRGFVGCKSRLQWAVSGTGL ASPLANANGTVFGCFAKKNVCDSKTCHNGGTC VQWDAFSCSCLIGFGCKSCAQBMAPQHF LGSSLVAWHGLSLPSQWYLSLMFRTRQAD GALATGRSSTITLQREGHWANSYGTGL QASSLRLEFGRANDGDWHHAQLALGAIGGF GGALLSFDYQGARAGGNLOPRHAHLHSNITV GGPGPAGGVARGFRGCLQGVRVSDTTEGVN SLDFSHGSINVSQGCSLPPPCDSNFCPANSY CSNDWDSYSCSCDPGSLPPCDSNFCPANSY CSNDWDSYSCSCDPGSLPPCDCSNFCPANSY CSNDWDSYSCSCDPGVSGCDCRONDFP BHQSVCTRKPSAHPGYTCCPFPNLOPYCET RIDQFCPRGWWGHPTCGPCNCDVSKGFDPDC NTSGEGCHCKENHTYRPSSPTCLLCCPYTG SLSRVCDFEDGGCCCKGGUGRCDCNCDPF SLSRVCDFEDGGCCCKGGUGRCGDCCDNFP AVTINGENNYDSGCSLPPCDSNFCPNSY CSNDWDSYSCSCDPGSALGRIWPRTRFG LPAAAPCFKGSFGTAVRHODEHRGWLPPLLF PRAAPCFKGSFGTAVRHODEHRGWLPPLLF NCTSITTSELKGFARGTRRESGLDSGRSQQL ALLLRNATHTAGYFGSDVKAYQLATELL AEESTGREGGCSATOWHTESHALTVSALL AEESTGREGGCSATOWHTESHALTVSALL AEESTGREGGCSATOWHTESHALTVSALL DTANKRIWELQOTEGGTAVLLOHYAAYAS ALAQMMRTILSFTLYTPNINSVYLVSALL AGENTRASHVELAGGEGGCCCCOPTEXAGGGCCCCCOPT CEPTAVRAGGGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC					sequence		
VFVEENSPIGLAVARVTATIPDEGTHAQMY (VVEGNIPVEVQLUISTGELTALVDLUSTERE YVLVIQATSAFLVSRATVHIVALLDRINDIPPY LONFELLENNYVTINESSFEGGAGESVALLASTGELKLSR ALDNRPLEAMSVLVSDGVHSVTACACALRV TITTDEMLTHSTITLELDMSPERFLSPLLGLFIQ AVAATLATPPDHVVYFWQRUTDARGGHILM VSLSVGQPPGPGGGPFLYSEDLGER; VLINRS LILTAISAGRYLPEDDNICLERFCENYMECVSV LRFDSSAFFLASSVLPRPIPVGGLRCCPPGF TODYCETEVDLCYSRPGPHGGRCSREGGYT CVNLLVGGFRCCDCSGDFERFYCQVTTRSFP ANSFITTRGLRQRHFTLALSFATKEDGLLL YNGRPNEKHDFVALEVIQEQVQLTFSAGEST TVSSFYPGGVSDGQWHTVQLKYYNKPLLG QTGLPQGPSEQKVAVVTVDCCTGVALRFGS VGNTSCAAQGTTGGSKKSLDLTGFLLLGG VPDLPESFFVBRRCFVGCWRNLQVDSRHDM APFIANNGTVPGCFAKINCVGSKTCHGGTC VNQWDAFSCEPLGFGGKSCAQEMAPPQHF LOSSLVAWHGLSLFSQPWYLSLMFRTRQAD GVLLQATTEGRSTTILQLRSGHVMLSVSGTGLB QASSLREFGRANDGDWHAQLALGAGGP GHALLSPDVGQCRAEGALGPRH-HGLHLSNTTV GGPGFAGGVARGFGCLGGVAVGDTFEGVN SLDPSHGESINVEQGCSLFDPCTSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC BRQQCPRGWWGHPTCQPCNCDVSKGDDPDC NRTSGEGCHCKENTYRPSSPTCLLOCYPTC SLSRVCDPEDGQCCCRPGVIGRCCDRCDNFF AVTITNCEVNTDSCFRALAGGTWWFTLFGF LPAAAPCRGSFGTAVRHODEHRGWLPPHLF NTSITTSELKGFARELYRSSPTSLOCYPTC SLSRVCDPEDGQCCCRPGVIGRCGDRCDNFF AVTITNCEVNTDSCFRALAGGTWWFTLFGF LPAAAPCRGSFGTAVRHODEHRGWLPPHLF NTSITTSELKGFARELQRRSEGGLBGSRQQL ALLLRNATQHTAGYFGSDVKAVAQLATRLL ARESTGRGGGLSATOQHTFEILNR VASALL ALGRREFSHELQGTEGGTARALQHVENGSLL DTANKRHWELQGTEGGTAWLQHVENGSL GGAVASVITYRTLAGLPHYYDDRSLRVPK BETPVVVRAGGGGGAGPELARGCHENGSAGQ ALLLRNATQHTAGYFGSDVKAVAQLATRLL ARESTGRGGGLSATOQHTFEILNR VASALL ALGRREFSHSSCCOCHMITSFALLBUNSRER NGEILPLKTLTYVALGVTLAGLLLTFFLTLL RILEBRENDRYCVSHAGGGGGRACCAGRAGGGGGRAGG EVYFRABSHSVCCCURMITSFALLBUNSRER NGEILPLKTLTYVALGVTLAGLLLTFFTTIL RILEBRENDRYCVSHAGGGGGGGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGG				1			
QIVEGNIPEVFQLDIPSGELTAL.VDLDYEDRYP LGNFELFNNYVTNSSSFPGGAIGRVPAIDD DISDSLTYSFERGELSLULIANSTGELKLSR ALDNNRPLEAMSVLYSDGVISVTAQCALK WITTDEMLTHSTILALEDMSFERTS.PLIGLIFQ AVAATLATPOHVVVNVQRDTDAPGGHIIN VSLSVQOPPGFGGGPFLSPLIARS LLTAISAQRVLPFDDNICLEFCENYARCVSV LRDSSAPFLASSVLFPIPHPVGCRYCNGGT CLCRDOYTGEHEVSARSGGCTPCYCCT TGDYCETEVDLCYSRPCOPHGRCSSREGGYT CLCRDOYTGHEVSVASSGGCTPCVCHOGT CVALLVGGRCDCSGDFERPYCQVTTRSFP AHSFITRGLRQRFHFTLALSFATKEEDGILL YNGRREHDEVALEVIGGQVGLTFSAGEST TTVSFPVFGVSDGGWHTVQLKYYNKPLL QTGLPQGFSGRVAVVTVDGCDTVALKFGS VLONYSCAAQGTQGGSKSLDLTGFLLGG VEDLPESFFVRMQFVGCMRALQDSSHDM ADPIANNGTVFGCFAKKNVCDSKTCHNGGTC VNQWDAFSCECTLOFGGKSCAGMANPQHF LGSSLVAWHGLSLPISQPWVLSLMFRTRQAD GVLLQATTRGRSTITLQLEGGHVMLSVEGTGL QASSLRLEFGRANDGDWHIAQLALAGIGGG GHALLSDYGQGAREONLGPRLIGLELSNITV GGPGPAGGVARGFRGCLQGVRVDTPECVY SLDFSHGESNFVGQCSLPPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLING GGPGPAGGVARGFRGCLQGVRVDTPECVY SLDFSHGSNNVGGCSLPPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLING GGPCPAGGVARGFRGCLQGVRVDTPECVY SLDFSHGSNNVGGCSLPPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLING GGPCPAGGVARGFRGCLQGVRVDTPECVY SLSFGCHCKENTYPROSPTCLCDCYPTC SLSRVCDPEGGCCCLPPCCDSNCCACDPPLC NKTSGECHCKENTYPROSPTCLCDCYPTC SLSRVCDPEGGCPCKPGVIGRQCDRODPP AEVTTNGCEVNYDSCPRALEAGHRGWARTHFF AFTNGCEVNYDSCPRALEAGHRGWARTHFF AFTNGCEVNYDSCPRALEAGHRGWARTHFF AFTNGCEVNYDSCPRALEAGHRGWARTHFF AFTNGCEVNYDSCPRALEAGHRGWARTHFF AFTNGCEVNYDSCPRALEAGHRGWARTHFF AGAALPRYALRGGOPPDLETTVILESSYPR ETPPVYNAPGCGGAGPEDLEARRGWARG LPAAAPCKGFCTAWHISILAGFREDSAGQ ALAQNMRHTYLSPTTTPTNYTVSLAKGAL ALAGNMRHTYLSFTTTPTTVTVSALLAGAL ALAGNMRHTYLSFTTTPTTVTVSALLAGAL ALAGNMRHTYLSFTTTTPTTVTTVSALLAGAL ALAGNMRHTYLSFTTTTTTTTTTLAGGWARG EVYFRASSFYSCQCHMTSFAVALLAGLHLLY RALIFERENTYCCYFWNSILLAGFGGWARGG EVYFRASSFYSCQCHMTSFAVALLAGLHLLY RALIFERENTYCCYFWNSILLAGTGGWARGG EVYFRASSFYSCGCHTLASTACACRGCFFKKG GGAVASVITYTLAGLLGFTCTTLL RILBSROHGGRNLTAALGLAQLVFTLIGFSTAD ALACCHREMELSACHARGCOPPECLANTCOSTYPAD ALACCHREMELSACHARGCOPPECLANTCOSTYPTLUSFAGP VAFAVSMSVLYTLAARASCAAQRCGFFKKG EVAFAVSMSVLYTLAARASCAAQRCGFFKKG EVAFAVSMSVLYTLAARASCAAQRCGFFKKG EVAFAVSMSVLYTLAARASCA	1						
YVLVIQATSAPI.VSRAT/HVRLLDRINDIPPI LGRFELFINYVTYNSSSFFGGAIGRYPAIPP DISDSLTYSFERGRELSLVLINASTGELKIS, ALDNINPLEAMBYLVSDGVISVTAQCALRV TIITDEMLTHSITI.RLEDMSFERT.SPLIGIT, AVAATLATPPDHVVTNVRQDTDAPGGHIN VSLSVQOPPGPGGGPPFI PSEDLQERLYLARS LLTAISAQRVLPPDDNICLEPCENYMRCVSV LRFDSSAPFIASSSVLFPPHIPVGGLRCCPPGF TGDYCETEVDLCYSRPCOPHGRCCSREGGYT CLCRDOYTGEHECVSARSGCTFOVCKNGGT CVNLLVGGFKCDCYSGRFGFTOVCKNGGT CVNLLVGGFKCDCYSGRFGRYQCYTTRSF AKSFITTRGLRGRFHTIALSF ATKERDGLLL YNGRNEKHDFVALEVIGEQVQLTFSAGEST TTVSFYPGGVSBQGWHTVQLKYNRELLG QTGLDQGFSQKVAVVTVDGCDTGVALFRGS VLGNYSCAAVGFGGGSKKSLDLTGFLLLGG VPDLPESFYRMRQFVGGMRNLQVDSRHIDM ADPIANNGTVPGGTGSKKSLDLTGFLLLGG VPDLPESFYRMRQFVGGMRNLQVDSRHIDM ADPIANNGTVPGGTAKKNVCDSKTCRINGGTC VNQWDAFSCECPLGFGGKSCAQEMANPOHF LGSSLVAWHGLSJFSOPWYLSKTCRINGGTC VNQWDAFSCECPLGFGGKSCAQEMANPOHF LGSSLVAWHGLSJFSOPWYLSKTCRINGGTC QASSLREPGRANDDOWHHAQLALGAIGGP GHALSFDYGOQRAEGNLGPRLHGLHLSNTTY GGPCPAGGVARGFGCLGGVAVDTPEGVY SLDPSHGESINVEGGCSLPPCDSNPCPANSY CSDWDSYSCSCOPGYGDNCTVCLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPVCET RDQPCFRGWWGHTCGFCNCDVSKGFDED NKTSGECHCKENTYRPGSPTLCLACCYPTIC SLSRVCDPEDGQCFCKPGVIGRQCDROLDNFF AEVITNGCEWNYDSCPRAEAGRWWFRTRFG LPAAAPCPKGSFGTAWRLCDEHRGWLPFRILL NCTSITFSELKGFAERLQRNESGLGSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATALL AHESTGRGFGLSATQDVHFTENLAUGHSAYAA ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTTVPLINGSMLXSAAL ALAQNMRHTLSFTTTTVPLINGSMLXSAAL ALAGNMRHTLSFTTTTTVPLINGSMLXSAAL ALAGNMRHTLSFTTTTTVPLINGSMLXSAAL ALAGNMRHTLSFTTTTTVPLINGSMLXSAAL ALAGNMRHTLSFTTTTTVPLINGSMLXSAAL ALAGNMRHTLSFTTTTTTVTLAGRACHELGRUPPTLAUGHTLAGL ALELBRACHARDAGGOPPELARTCYNCLSFYADG GEAVASVITYTLAGLGCPPTDLTVTLLGSVFT ANGELPLATLTTVALLGVTLAALLLTFFFTTT ALAGLEPGGCPTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT							
LGNFELENVTNRSSSPGAGARRYABID DISDSLTYSFERGELSLULNSTGELKISR ALDNNRPLEAMSVLYSDGVISVTAQCALISR ALDNNRPLEAMSVLYSDGVISVTAQCALISR TIITDEMLTHSTILLEMSSFERT.SPLIGLEQUAVAATLATPPHVVVPNVRDTDAFOGHILN VSLSVQOPPGOGGPPFLPSEDLQERLYLTNS LLTAISAQRVLPFDDNICLRPECENYMRCVSV LRPDSSAPFIASSSVLFRPIHPVGGLRCCPPGF TGDYCGTEVDLVSRFCCPHAGRCSREGGYT CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGEHCEVSARSGRCTPGVCKNOGG CLCRDGYTGHCHCSTPGVCKNOGG CLCRDGYTGHCAGVACAUTUGHCHCSTPG AUSTAINATUGHCHCAGATAGAGG GYLLQATTGRGTTTLQLREGGTVALSVEGTGI QASSLALEFGRANDGWHCHGLHLSTNTTV GURGAAGGVARGTAGAGAGGG GHALSFDVQQCRAEGNLGPRLHGLHLSTNTTV GURGAAGGVARGTAGCAGGYRVSDTPEGVN SLDPSHGESNVEGGCSLPPCDSNPCPANSY CSDDWDSVSCSCDFGYCGNOTPCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPVCET RDQPCRGWWGHPTGCPNCDDNSGFDPD NKTSGECHCKNNYRPGSPTCLLCDCYPTG SLSRVCDPEGGGCPCKPGVIGRGTCDCNDPP AEVTINGCCNYDSCPRABAGWWRTTEFG LPAAAPCPKGSGFAVARICDEHPTFLRVGSALL DTANKGHPELGGTAWLLQHFLAYASA ALAQMRHTYLSFTITYPHVISVYRLDKGN FAGALLRYBALRGQPPDLETTVLLESSYR PETPVVVRAGPGGCPERLARRGRHPELSQ GEAVASVITYRTLAGLLPHNYDPDKRSLRYPY RIENTPVVSSVPDEELLPRALLBCHSUL ALLERESTQRGCJSATQUHTFILLBCGGAAG LLTYPLYBALGGCPPTLLRYWDPDKRSLRYPY RIENTPVVSSVPDEELLPRALLBCHJUL RIESSTQRGCJSATQUHTFILLBCGGAWARGC EVYFRNSHYSCGCNHMTSFAVALUDARRG NGELPLKTLTYVALGVTLAALLLTFFFT.LL RIESSTQRGCJSATQUHTFILLBCGGAWARGC EVYFRNSHYSCGCCHMTSFAVALUDARRG NGELPLKTLTYVALGVTLAALLLTFFFTLLL RIESSTQRGCPSFAVLLLLALLLTFFTLLL RIESSTQRGCPFBT-LSYVLYRDVSRAGC EVYFRNSHYSCGCCHMTSFAVALUDARRG NGELPLKTLTYVALGVTLAALLLTFFFTLLL RIESSTQRGCPFBT-LSYVLYRDVSRAGC EVYFRNSHYSCGCCHMTSFAVALUDARRG NGELPLKTLTYVALGVTLAALLLTFFFTLL LAVGLOPPGGRPTY-LLSTANDVSRCGC							
DISSILTYSFERGNELSULINASTIGELIS. ALDNINPLEAMBYU-YSIGVISYTAQCALRY TIITDEMLTHSITLR.LEDMSFERT.SPILGIE, AVAATLATPPHVVFNYQRDTDAFOGHIIN VSLSVGQPPGPGGGPPFLSFELGER.YYLRYS LLTAISAQRVLPFDDNICLREPCENYMRCYSY LRDSSAPFIASSSVLFRIHPYGGLRCRCPPGF TGDYCETEVDLCYSRPCCPHGRCRSERGGYT CLCRDOYTGEHECVSARSGRCTPOVCKNGGT CVALLVGGFKCDCPSGDFERPYQCYTTSFF AKSFITTRGIR.GRFHFTLALSFARKERDGLLL YNGRPNEKHDFVALEVIGGQVQLTFSAGEST TYSSPYPGGVSDGOWHTYQLKYYNKPLLG QTGLPQGPSGKVAVVTVDGCDTGVALFFGS VLGNYSCAAVGTOGGSKKSLDLTGFLLGG VPDLPSFYPRRGVFOGMSKSLDLTGFLLGG VPDLPSFYPRRGVFOGMSKSLDLTGFLLGG VPDLPSFYPRRGVFOGMSKSLDLTGFLLGG VPDLPSFYPRRGVFOGMSKSLDLTGFLLGG VPDLPSFYPRRGVFOCMRALQVDSRHIDM ADPIANNGTVPGCFAKKNVCDSKTCHNGGTC VNQWDAFSCELFDFGGKSCAGMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRITQAA GVLLQATIRGRSTITLQLREGHTVMLSVEGTGL QASSLRLEPGRANDGDWHIAQLALGAIGGF GHALISPDYQQGARGONLGPRLHGLBLSNITV GGPGPAGGVARGFRGCLQGVRVDTPEGVON SLDPSHGESINVGGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHGSVCTRKESAPHGYTCECPPNYLOPLYCET RDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSFTCLLCDCYPTI SLSRVCOTRESAPHGYTCECPPNYLOPLYCET RDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSFTCLLCDCYPTI SLSRVCOTRECAPHGYTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSFTCLLCDCYPTI SLSRVCOTRECAPHGYTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPFGSFTCLLCDCYPTI SLSRVCOTREGAPHGYTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPFGSFTCLLCDCYPTI SLSRVCOTREGAPHGYTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPFGSFTCLLCDCYPTIF SLSRVCOTREGAPHGYTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPFGSFTCLLCDCYPTIF SLSRVCOTREGAPHGTAVHTCDEHREAUPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQOL ALLINATQHTHAGYGSDVKVAQUATRLL AHESTORGFGLSATQDVFITFENLLRVGSALL DIANKRHWELIQOTGGTMAULDLEPSFFP ETPPVVRPAGPGEAQEPEELARRQRAPPELSQ GGAVASTUTRTLAGLPRHDVPDFKSLKTYP RIPMTPVVSISVHDDEELLPRALDKPTTVGPF LLETERTKRPICYFWNNSILVSGTGGWSARGC EVVFRNSHNSCQCONHMTSTAVLTDENSFYADG GGAVASTUTRTLAALLLATFFFTLTLESVFY RIPMTPVVSISVHDDEELLPRALDKPTTVGFF LAVGLDPEGGGMPATPTTI LAVGLDPEGGGPPATTTTAALGLAGLPTLAUDVSRRE NGELLPLKTLTYVALGVTLAALLLTFFFTLLL RLESNOHGFRFTVLALGGOPPDLETTSVVLSKEVKRAL LAVGLOPPGGFGFAKG PVSGLOPSFAVILLLAATTAALGLAGVFTTTOR VAFAYSMSVFLYTLAAALLAGROFFTTTOR LA							
ALDNNRPLEAMSVLYSDCVISVTAQCALIST VITIDEMLTHSTIRLEDMSPERTS.PLIGIFQ AVAATLATPOHVVPNVQRDTDAPGGHIND VSLSVQOPPGGGGGPPET.PSEID.GRIVJANS LLTAISAQRVLPFDDNICLREPCENTYMRCVSV LRPDSSAPFIASSSVLFRPHIPVGGLRCCPPGF TGDYCGTEVDLCVSRPCOPHAGCGSREGGYT CLCRDOYTGEHCEVSARSGRCTPGVCKNGGGT CVALL VGGFKCDCYSRPCOPHAGCGSREGGYT CLCRDOYTGEHCEVSARSGRCTPGVCKNGGT CVALL VGGFKCDCSGDFEKFYCQVTTRSFP AHSFITEGLRQRFHFTLALSFATKERGGLL YNGRYBEKHDEVALEVIQEQVDLTSAGEST TTVSFFVFGGVSBGQWHTVQLKYYMKPLLG QTGLPQGFSgCWAVVTVDGGTVALFGGSSVLGNTSAGEST TTVSFFVFGGVSBGQWHTVQLKYYMKPLLG QTGLPQGFSGCWAVVTVTDGCTVALFGGS VLGNTSFCAAQGTOGGSKKSLDLTGFLLLGGG VPDLPESFPVRMGFVGCMRNLQVDSRHDM ADPIANNGTVPGCFAKKNVCDSKTCHNGGG VPDLPESFPVRMGFVGCMRNLQVDSRHDM ADPIANNGTVPGCFAKKNVCDSKTCHNGGG QVDLPESFPVRMGFVGCMRNLQVDSRHDM ADPIANNGTVPGCFAKKNVCDSKTCHNGGG QVDLPESFPVRMGFVGCMRNLGVBSRHDM ADPIANNGTVPGCFAKKNVCDSKTCHNGGG QVDLPESFPVRMGFVGCMRNLGVBSRHDM ADPIANNGTVPGCFAKKNVCDSKTCHNGGG QVDLPESFPVRMGFVGCMRNLGVBSRHDM ADPIANNGTVPGCFAKKNVCDSKTCHNGGG QVDLPESFPVRMGFVGCGNAMPOHF LGSSLVAWHGLSLPISQPWLSLMFTRQAD GVILQAITRGRTTTLQLREGHVMLSVEGTOL QASSLRLEPGRANDGDWHLAGLALGAIGGF GHALLSFDVQQGRAEGNLGFRLGLGLANTUTV GGIPGAGGVARGFRGCLQGVYSDTPEGVN SLDPSHGESNNVCGGCSLDPCDSNPCPANSY CSNDWDSVSSCDFGYGDNCTNVCDLNPC BHQSVCTKKPSAPHGYTCECPPNYLGPVCET RDQCPCRGWWIGHTTCGPCCDVSKGFDPDC NKTSGGCHCKENHYRPGSFTCLLCDCYTTG SLSKVCDPEGGCPCKGPVIGRGCDRCDNPF AEVTTNGCCNYDDSCPRALBGGWWRTTEFG LSAACCAGCCFCKAPVIGRCCPCCPCAFVIGRCCPCTCHNTSPAACAACAACCAGCFKGCAGCGCCCNCNPF AEVTTNGCCNYDDSCPRALBGGWWRTTEFG LDAAAPCKGSGTAVRTCDEHGGWPPHLF NCTSITSELKGFABLQRBGGLDSGRSQOL ALLLENATQHTAGYGSDVXAYQLATRILL AHESTQRGGGLSATQAWCHTELKVGSAL DTANKHWBLQQTEGGTAWLLQHYEAYASAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA							l
THITDEM_THSTILR_LEDMSPERFI.SPL_LGLF_IQ AVAATLATPOPHVVYENVORDAPGGHIIN VSI_SVGOPPGGGGPPLPSEDL_GER_YLNRS LLTAISAQRV_IPEDDNICL_RESVAYRCVYS LRFDSSAPFIASSSV_FRPHPVGGLRCRCPPGF TGDYCETEVDL_CYSRPCOPHGRCRSREGGYT CLCRDGYTGEHGEVSARS.RGTAVCKNGGT CVNLLVGGFKCDCPSGDFEKPYCQVTITSSPF AHSFITTRGL_RGYHFITLALSFAYEREDGLLL YNGRPEKHDFVALEVUGGVQ_UTTSAGEST TTVSPFVFGGVSDQWHTVQCYTTSSPF AHSFITTRGL_RGYHFITLALSFATKERDGLLL GTGL_PQGPSEQ_KVAVVTVDGCDTGVALRFGS VLGMYSCAAQGTGGGSK_KYYNYRLL_G GTGL_PQGPSEQ_KVAVVTVDGCDTGVALRFGS VLGMYSCAAQGTGGGSK_KAYNYNKPLL_G GTGL_PQGPSEQ_KVAVVTVDGCDTGVALRFGS VLGMYSCAAQGTGGGSK_SCAQEMANNPOHF LGSSLVAWHGLSLPISOPWYLSLAMTRQ_AD GVLLQAITRGRSTITLQ_IREGHL_SHITV GROWDAPSCECPL_GFGGKSCAQEMANNPOHF LGSSLVAWHGLSLPISOPWYLSLAMTRQ_AD GVLLQAITRGRSTITLQ_IREGHL_GLL_SHITV GGMCPAGGVARGFRGCL_QGVRVSDTPBCVN SLDPSHGSSINVEQGCS_UPPCDSNPCPANSY CSNDWDSYSCSCDPGYQDNCTNNCDLNPC GHGSATSPTYGQGQAAGGTGCL_QGVRVSDTPBCVN SLDPSHGSSINVEQGCS_UPPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNNCDLNPC EHQSVCTRYPASPHGYTCECPPNYL_GPYCET RDQPCPRGWWGHPTOGPCNCDVSKGFDPDC NKTSGECHCKENHYRPQSFTCL_LCDCYPTOF SLSRYCDPEDGQCPCKPGYIGRQCDRCDNPF AEVTTNGCEVNYDSCPRALEAGIWWFRIFG JEAAAPCKRGSGTAVRACDERRGWLPPNLF, NCTSITTSELKGFABRL_QRNESGLDSGRSQQL ALLLRANTQHTAGSYTGSDVXAYQLATRLL AHBSTGRGFGL_SATQDVHFTENLLRVGSALL DTANKRHWELLQOTTGGTAWLDPYLAYAS ALAQNMRHTYLSFTTYTPTNVISVVAYQLATRLL AHBSTGRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELLQOTTGGTAWLDPYLAYAS ALAQNMRHTYLSFTTYTPTNVISVVAYQLATRLL AHBSTGRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELLQOTTGGTAWCLPYPLYFEY ETFPVVRPAGPGEAQEPELARQRRHPELSQ GGAVASVINYTLAGLICHTTVILPESVPR ETFPVVRPAGPGEAQEPELARQRRHPELSQ GGAVASVINYTLAGLICHTTVILPESVPR ETFPVVRPAGPGEAQEPELARQRRHPELSQ GGAVASVINYTLAGLICHTTVILPESVPR ETFPVVRPAGPGGAQEPELARQRRHPELSQ GGAVASVINYTLAGLICHTTVILPESVPR RITHTPVVSINYDDEBLLFRALDKPVTVQFR LLETEERRECVFWMHSILVSGTGGWSAGC EVVFRNESHVSQCQNHMTSTALJLKVFR RITHTPVSINYTHALARLSCAAQRQGFEKKG PVSGLQPSFAVLLLSATWLLALLLNYDTILL HERLERNOHGFRRATLALAGLYCHLGRKDQ VAFAYMSWSTYLTLAARASCAAQRQGFEKKG PVSGLQPSFAVLLLSATWLLALILSNNDTILL HTYLPATOTIOQPFFLSYVVLSKEVKALKA LACSKRSPDPQFGLGGGFGGACGFCGRKDQ VAFAYMSMSTL							
AVAATLATPPDHVVFNQRDTDAGGHIIN VSLSVQOPPGGGFPLYSELDLGERLYLNRS LLTAISAQRVLPFDDNICLREPCENYARCVSV LRFDSSAPFLASSVLFRPHPVGGLACRCPPGF TGDYCETEVDLCYSRPCGPHGRCRSREGGYT CLCRDGYTGEHCEVSARSGRCTFGVCKNOGT CVNLLVGFRCKCDCSGDFERFYQCVTTRSFP AHSFITTRGLRQRFHFTLALSFATKERDGLL YNGFFNEKHDFVALEVIGGOVQLTPSAGEST TTVSPFVGGVSDGQWHTVQLKYYNFLLG QTGLPQGPSEQKVAVVTVDFGCDTGVALRFGS VLGNYSCAA\QGTQGGSKKSLDLTGPLLGG VPDLF2SFPVRMRQFVGCMNLQVDSRHDM ADFIANNGTVFGCFAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPOHF LGSSLVAWHGLSLFISQPWYLSLMFRTRQAD GVLLQAITRGRSTITLQLREGHVMLSVFGTGL QASSIRLEPGRANDGDWHHAQLALGAIGGF GHALSFDYGQQRAEGRLGPRLHGLHLSNTV GGPGPAGGVARGFRGLQCYSDTPECVN SLDPSHGESINVEQGCSLFDFCDSNPCPANSY CSNDWBSYSCSCDPGYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPNVLSCPFPNVLGFPCE RDQCPGRGWWGHPTCGFCNCDVSKGFDPC NKTSGECHCKEMHYXPPGSTCLLCCDCYPTG SLSRVCDPEDGQCPCKFGVIGRQCDRCDNPF AEVTTNGCEVNYDSCPRAEAGIWWFRTFGG LPAAAPCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGWWFRTFGG LPAAAPCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGWWFRTFGF LPAAAPCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFGF LPAAAPCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFGF LPAAAPCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAPCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFCFKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFAERLGYBWFRTFFG LPAAAFSTWALLAGLLEHNFULNGTUGALLALLTFFLILL AHESTGRGGLSATQDVHTTENLLKVGSALL DIAMRHWELLQOTTGGGTAALLQUFGTYGTAAALLLTFFLILL RLRSNQHGRGRKALAGUPVALGWFYAFTG LLETERTKPICVFWNHSILVGTTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDWSRE NGELPLKTLTTVALGUPLAALLLTFFFLILL RLRSNQHGRRNLTAALGLLARQGGFEKKA PVGLQPSFAVLLLLSATWLLALLSVNSDTLL FMYLTATONCIQOFFLSYVULSKEVKKALK LACSKRSPDPALTTKSTLTSSTYDCTSSPADE VAFAVSMSVFLTILAARASCAGCFEKKAG PVGLQPSFAVLLLLSATWSDTLL FMYLTATO							I ·
VSLSVGOPFGGGFFLESDLOGERLYNRS LLTAISAGRVLPFEDNICLREPCENYMRCVSV LRFDSSAPFIASSSVLFRPHFPVGGLRCRCPPGF TGDVCETEVDLCYSRPCOPHGRCRSREGGYT CLCRDGYTGHEGEVSARSGRCTFGVCKNGGT CVNLLVGGFKCDCFSGDFERFYCQVTITSFF AHSFITTRGLRGKFHFTLALSFATKERDGLLL YNGRFMEKHDFVALEVQEQVQLTFSAGEST TVSPFVFGGVSDGQWHTVQKYTNKFLLG QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS VLGMYSCAAQGTQGGSKSLQVYNSKPLLG QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS VLGMYSCAAQGTQGGSKSLDLTGFLLLGG VPDLPSSFPVRMRGPVGCMRNLQVDSKHDM ADFIANNGTVPGCTAKKNVCDSKTCHINGGTC VNQWDAFSCSCPLGFGGSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLAFRTRQAD GVLLQATTRGRSTITLQLREGHVMLSVEGTGL QASSURLEPGRANDGDWHHAQLALGAIGGF GHALISFDYQGQAEGNLGPRLHGHLSNITV GGPGPAGGVARGFRGGLQGVRVSDTPEGVN SLDPSHGSSINVEQGCSLPPDCSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RDQPCPRGWWGHPTCGPCNCDVSKGFDPD NKTSGECHCKEHYTPPOSTLCDCYPTG SLSRYVCDPEDGQCPCKPGVIGRQCDRCDNFF AEVTTNGCEVNTDSCPRALEAGIWWFRTRFG LPAAAPCPKGSFGTAVHCDEHRGWLPPNLF NCTSITTSELKGFABRLQRNESGLDSGRSQQL ALLLRANGTHAGYTGSDVXAVQJATRLL AHESTGRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELLGOTTEGGTANGTHYAGSVALL AHESTGRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELLGOTTEGGTANGTHYAGSVALL LATESTAVLTSVALGGGEPELARRQRREFELSQ GEAVASVIIYRTLAGLICPTTTULPESVPR ETPPVVRAGPGEAQEPELARRQRREFELSQ GEAVASVIIYRTLAGLLLFHTYTLESVRR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVVSISVHDDEELLFRALDKPVTVQFR RPINTPVSISVHDDEELLFRALDKPVTVQFR RPINTPVSISVHDDEELLFRALDKPVTVQFR RPINTPVSISVHDDEELLFRALDKPVTVQFR RPINTPVSISVHDEHLSAAAQGRGEKKG PVSGLQPSFAVLLLLSATWLLALLLSVNSDTIL RALTERDDVTTGPRAPTYMGACAAQGREKKALK LACSKRSPDPALTIKSTLTSSCKSQSPSYP							
LLTAISAQRVLPFDDNICLREPCENTMRCVSV LRPDSSAPPLASSVLRPHEVGGLRCRCPDGF TGDVCETEVDLCYSRPCOPHGRCRSEGGYT CLCRDGYTGEHCEVSARSGRCTPQVCKNGGT CVNLLVGGFKCDCPSGDFERPYQVCKNGGT CVNLLVGGFKCDCPSGDFERPYQVCTTRSFP AHSFITTRGLRQRFHFTLALSFATKERGLLL YNGFFMEKHDFVALEVUGGUTFSAGEST TTVSPFVPGGVSDQWHTVQLKYYNKPLLG QTGLPQOPSEGVXAVVTVDGCDTGVALRFGS VLGNYSCAAQGTQGGSKSLDLTGPLLLGG VPDLPSSFPVBMRGFVCGMKRNLQVDSRHDM ADFLANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEWANNOHF LGSSLVAWHGLSLPSQPWYLSLMFRTRQAD GVLLQATRGRSTITLQLREGHVMLSVEGTGL QASSLLPSGRANDGDWHHAQLALGAIGGF GHALSFDYGQQRAEGNLGGRLHGHLSNITV GGPGPAGGVAGRFGCLQCAYSDTPEGVN SLDPSHIGSSINVEQGCSLPPPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCCFPNYLGPYCET RDQCPRGWWGHPTCGFCNCDVSKGFDDC NKTSGECHCKEMHYXPPGSTCLLCDCYPTG SLSRVCDPEDQQCCKFGVIGRQCDRCDNPF AEVTTNGCEVNYDSCPRAEGWWPRTTFG LPAAAPCPKGSFGTAVHCDEHRGWLPPHLF NCTSITTSELKGPAERQWFREGUFGGDGSGNQQL ALLLRNATGHTAGYTEGDVXVAYQLATRLL AHESTQRGFGLATQDVHFTENLLRVGSALL DTANKRHWELQQTEGGTAWLLQHYBAYAS ALAQNMRHTYLSFTITVTPNIVISVVRLDKGN FAGAKLPRVARGGEAQEPEELARQRRHPELSQ GEAVASVINTALAGLLEHNYDPOKSLRVPK RPINTPVVSISVHDDEELLPRALDKPVTVQFR LLETERTKPICTYWHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGELPLKTLTTVALGUTLAALLLTFFTLLL RILBSNOHGIRRNITAAGLAQLVFLLGNGWSAGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGELPLKTLTTVALGUTLAALLLTFTLLL RILBSNOHGIRRNITAAGLAQLVFLLGNGWSAGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGELPLKTLTTVALGVTLAALLLTSTLLL RILBSNOHGIRRNITAAGLAQLVFLLGNGWSAGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGELPLKTLTTVALGUTLAALLLSVNSDTLL FHYLFATCNCIQGPFKAGAC VAFAVSWISTYLTLAARLSACNGGFEKKAG VAFAVSWISTYLTLAARLSACNGGFSKAG VAFAVSWISTYLTLAARLSACNGGFSKAG VAFAVSWISTYLTLAARLSACNGGFSKAG PVSGLOPSFAVLLLALSTWILLALLSVNSDTLL FHYLFATCNCIQGPFILSYVVLSKEVKKALK LACSRKSFDPPALITKSTLTSSYCLSFYAGD VAFAVSWISTYLTLAARLSCNCFSFYADG RACSRRSFDPPALITKSTLTSSTGRSCKSOPSYPPLIR ESSALNFGQGPFELGFGGRFKDQG VAFAVSWISTYLTLAARLSCNCFSFYADG RACSRRSFDPPALITKSTLTSSTGRSCKSOPSYPPLIR EESALNFGQGPFGLGGFGGRACCFLGGRFKDQG VAFAVSMSVSTLTLAARLSCRCFSFYADG RACSRRSFDPPALITKSTLTSSTGRSCKSOPSYPTPLIR EESALNFGQGPFGLAGGFGCKGGR							
LRFDSSAPFLASSSVLFRPIHPVGGLRCRCPPGE TGDYCETEVDLCVSRRCPHGRCRSREGGYT CLCXDLYGGFRCDCCSPGHGRCRSREGGYT CLCXDLYGGFRCDCCSGDFRCPYCCVTRSFF AHSTITRGLRQRHFTLALSFATKERDGLLL YNGRRNEKHDFVALEVIQEQUQLTFSAGEST TTVSFFVFGGVSDGQWHYCLKYYNKPLLG QTGLPQGPSEQKVAVVTVDGCDTOVALRFGS VLGAYSCAAQGTQGGSKSLDLTGPLLLGG VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM ADFIANNGTVPGCFAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQATIRGRSTITLQLRGMMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGOF GHALLSFDVQQRAEGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGOF GHALLSFDVQQRAEGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGOF GHALLSFDVGQCRAEGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGOF GHALSFDVGQCRAEGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGOF GHALSFDVGQCSLPPDCDSNFCPANSY CSNDWDSYSVGSCDPGYTGGNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHTFCCPCNCDVSKGFDPDC NKTSGECHCKENHYRPGSPTCLLCDCYPTG SLSRYCDPEDGQCPCKEPGRQCDCRODNPP AEVTTNGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLFN NCTSHTSELKGFAERLQRRESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTORRGGLSATDDWHALDLAKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGGLASTDDWTDVIVSVYRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGGAQEPELARRQRRIPELSQ GEAVASVIIYRTLAGLLPHYDPDKRSLRVPK RPINTTPVVSISVHDDEELLPRALDKVSALG GEAVASVIIYRTLAGLLPHYDPDKRSLRVPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVSISVHDDEELLPRALDKVSTLPK RPINTTPVVSISVHDDEELLPRALDKVSTLPK RPINTTPVVSISVHDDEELLPRALDKVSTLPK RPINTTPVVSISVHDDEELLPRALDKVSTLL LRIERSNQHGRRNLTAALGALVFTLIGHNQA DLFFACTVIALLHFLYLCTTSWALGHLAHLY RAALTEVRDVNTOPMRFYYMLGWGYPAFTIG LAVGLOPSGAVLLLSATWULLALLSVNISDTLL FHYLFACTOQOPFILSVVVLSKEVKALK LACSRYPSPDPALITTSSTLTSSYNCPSFYADG RXQPYYGDSAGSLHSTNSRGSKGYSPYPTLLR ESSALNGFQGOPFGLGGIFRLCGTLGRRKDQQ						1	
TGDYCETEVDLCYSRPCGPHGRCRSEGGYT CLCRONTGEHCEVSARSGRCTPGVCKNGGT CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP AHSFITFRGLLQRFHFTLALSFATKERDGLLL YNGRRNEKHDFVALEVUGQVQLTFSAGEST TTVSPFVPGGVSBQWHTVQLKYNKPLLG GTGLPQGPSEQXVAVYTVDGCDTGVALRFGS VLGNYSCAAQGTGGGSKXSLDLTGFLLLGG VFDLPESFFVRMRQFVGCMRNLQVDSRHDIM ADFIANNGTVPGCFAKKNVCDSKTCHNGGTC VNQWDAFSCECFLGFGGKSCAQEMANQHF LGSSLVAWHGLSLPISOPKISLMFRTRQAD GVLLQATTRGRSTITLQLREGHYMLSVEGTGL QASSLRLEPGANDGDWHAQLALGAIGGP GHAILSFDYQQRAEGNIAPQHAFGLESSEDVSQCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRRFSAPHGYTCEPPNVLGPYCET RDQPCPRGWWGHFTCCPCNCDVSKGFDPDC NKTSGECHCKENHYRPGSPTCLLCCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNFF AEVTTNGCENYDSCFRAEAGIWWRTTRFG LPAAAPCPKGSGTAVRGFGSDLVSVAYQLATRLL AHSSTGRGGSGTAVRHTENLRVGSALL DTANKRHWELLQQTEGGTAWLLQHYEAYAS ALAQNMRTTVJSFTTVTPINVISVVRLDKGN FAGAKLPRYEALRGCPPDLETTVILPESVFR ETPPVVRAPAGPGGAGDGTGAWLLQHYEAYAS ALAQNMRTTVJSFTTVTPINVISVVRLDKGN FAGAKLPRYEALRGCPPDLETTVILPESVFR ETPPVVRAPAGPGGAGDFGLARQRAFPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPINITPVVSISVHDDEELLFRALDKFVTVQFR LLETEERIKPLCVFWNNISLLVSGTGGWSARGC EVVFRNESHYSCQCNHMTSFAVLALLTFETLL RILSBRHGGPRDLAGLLFRALDKFVTVQFR LLETEERIKPLCVFWNNISLLVSGTGGWSARGC EVVFRNESHYSCQCNHMTSFAVLABLLTGHTVPL RALTEVRLVTVFMGWTFFFLTLL RILSBRQHGRRNLTAALLAGLAQLVFLLGINQA DLPFACTVIALLHFLYLTTSWALDHDVSREE NGELPLKTTTVALGVTLAALLLTFFTLTL RILSBRQHGRRNLTAALLAGLAQLVFLLGINQA DLPFACTVIALLHFLYLCTFSWALLAHLY RALTEVRDVTTGPMRFYYLLGGGPNSARGC EVVFRNESHYSCQCNHMTSFAVLABLLHDYSGRE UEGLPLKTTTVALGVTLAALLLTFTLTLL RILRBRQHGRRNLTAALLAGLAQLVFLLGINQA DLPFACTVIALLHFLYLCTFSWALLAHLY RALTEVRDVTTGPMRFYYLLGGGPNSARGC EVVFRNESHYSCQCNHMTSFAVLABLLHDYSGRE NGELPLKTTTVALGGVTAATITG LAVGLDPEGYGNPDFCWLSIYDTLLWSFAGP VARAVSMSVYLTLAARASCAAQRGGFEKKG PVSGLQPSFAVLLLLSATWLLALLLSVNSDTLL HAVLTEVRQOPPGLOUPSLKEVKKALK LACSRKPSPDPALITKSTLTSSYNCSPSYADG RXQPYSGLOPSGAULLLSATWLALLLESLENLQ ESSALNPGGQOPFGLGGTGGLCTLGRRKDQQ]						
CLCRDYTGEHCEVSARSORCTPGVCKNGGT CVNLLVGGFKCDCSGPEKPYCQVTTRSFP AHSFITFRGLRQRFHFTLALSFATKERDGLLL YNGRNEKHDFVALEVIQEVQCLITSAGEST TTVSFFVPGVSDGQWHTVQLKYYNKPLLG QTGLPQGPSEQKVAVVTVDGCDTQLTALFGGS VLGAYSCAAQGTGGGSKSLDLTGFLLLGG VPDLPESFPVRMRGFVGCMRNLQVDSRHDM ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQFWYLSLMFRTRQAD GVLLQATIRGRSTITIQLREGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGGF GHALSFDVGQRAEGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGGF GHALSFDVGQRAEGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGGF GHALSFDVGQRAEGHVMLSVEGTGL QASSLRLEFGRANDGDWHHAQLALGAIGGF GHALSFDVGQCSLPDFCDSNPCPANSY CSNDWDSYSCSCDPGYGDNCTNVCDLNPC EHQSVCTRKFSAPHGYTCECPPNTJGPYCET RIDQPCPRGWWGHTFCCPCNCDVSKGFDPDC NKTSGECHCKENHYRPGSFTCLLCCYPTG SLSRVCDPEDGQCPCKPGYGRQCDRCDNPF AEVITNGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDERWLPNLF NCTSHTFSELKGFAERLQRHSSGLDSGRSQQL ALLLRNATQHTAGYFGSDVKJAYQLATRLL AHESTGRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELLQOTEGGTAWLLQHYEAYAS ALAQNMRTHJSPFTUTPIVSVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRFAGPGEAQEPELLARRQRHPELSQ GEAVASVIIYRTLAGLLPHYDPDKRSLRVPK RPINITPVVSISVHDDEELLPRALDKFVTVQFF LLETEERTKFLCVFWNNSILVSGTGGWSARGC EVYFRNESHYSCQCHMTSFAVLMDWSRRE NGELPLKTLTYVALGWTTJUSFAGP VAFAVSNSVYLTILAARASCLAARRGGFEKKG PVSGLOPSFAVLLLLSATWLLALLLTFFFT-ILL RILRSNQHGRRNLTAALGLAQLVFLLGINQA DLPFACTVIALLHFLYLCTFSWALLLTHFFT-ILL RILRSNQHGRRNLTAALGLAQLVFLLGINQA DLPFACTVIALLHFLYLCTFSWALLLHLYFFALL LAYGLOPEGYGNPDCWLSIVDTLUSFAGP VAFAVSNSVYLVILARASCAAQRQGFEKKG PVSGLOPSFAVLLLLSATWVLALLLSVNSDTLL FHYLFACTACIQGPFIELSVVLSKEVRKALK LACSRKPSPDPALITKSTLTSSYNCPSFYADG RXQPYYGDSAGSLHSTNSRGSKGSYPYPTLLR ESSALNFGQOPPGLGFFLLGRRKKDQQ	.						
CVNLLVGGFKCDCPSGDPEKPYCQVTTRSFP ANSFITFRGIRQREPHTALISFATKERDGILL YNGRFNEKHDFVALEVIQEQVQLTPSAGEST TITVSFFVPGGVSDGQWHTVQLKYYNKPLLG QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS VLGNYSCAA\QGTQGGSKKSLDLTGFLLLGG QYDLPESFPVRMGVGCMRNLQVDSRHDM ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCCCPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTQAD GVLLQATTRGRSTITLQLREGHVMLSVEGTGL QASSLRLEPGRANDGWHAQLALGAIGGP GHALLSPDYQOQRAEGNLGPRLHGLHLSNITV GGPGPAGGVARGFRGCLQVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTIKKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVTINGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPPCKGSFGTAVRHCDEHGWH.PPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLKRNATOHTAGYBSOVKAYQLATRLL AHESTORGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLUHYEAYAS ALAQNMRHYLSPTTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETVILLPSVPR ETPPVVPRAGFGEAGEPELARRORRIPELSQ GEAVASVIJYRTLAGLLPHNYDPDKRSLRVPK RPINTPVVSISVHDDEELLPRALLDKPYTVGFR LLETEERTKPICVFWHNSILVSGTGGWSARGC EVVFRNSSHYSCQCNHMTSFAVLMDVSRRE MGELPLKTLTTVALGLAGLAGLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTSROHGRRNLTAALGLAGLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTSROHGRRNLTAALGLAGLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTSROHGRRNLTAALGLAGLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTSROHGPGRYCHGCPSKAYCHLLLSANNSDTILL FHYLFACTOCQOPPICLGGTGRKXGCPSVKPTPLLR LAVGLDPEGYGNPDFCWLSTYDTLIWSFAGP VAFAVSMSVFLYILLASACAAQRQGFEKKG PVSGLOPSFAVLLLLSATWLLALLSVNSDTILL FHYLFACTOCQOPPICLGGTGRKNCG FYGGLOPSFAVLLLLSATWLLALLSVNSDTILL FHYLFACTOCQOPPICLGGTGRKNCG							
AHSFITTRGLRQRFHFTLALSFATKERDGLLL YNGRPHEKHDFVALEQQUQLTFSAGEST TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG QTGLPQGFSEGKVAVVTVDGCDTGVALRFGS VLGNYSCAAVGGTGGGSKKSLDLTGPLLLGG VPDLPESFPVRMRQFVGCMRNLQVDSRHDM ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSMFRTRQAD GVLQAITGRGSTTILQLREGHVMLSVEGTGL QASSLRLEPGRANDGDWHHAQLALGAIGGP GHAILSFDYGQQRAEGNLGPRLHGHLSNITV GGIPGFAAGGVARGFRGLQGVRVSDTPEGVN SLDPSHGGSINPVEQGCSLPPDCDSNPCPANSY CSNDWDSYSCSCDPGYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPGSPTCLLCCYPTG SLSRVCDPEDQCPCKPGVIGRQCDRCDNFF AEVTINGCENYNDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTGATWLLQHYSAYAA ALAQNMRHTYLSFFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRHPELSQ GEAVASVIYRTLAGLLPHNYDDKRSLRVPK RPINTPVVSISVHDDEGPEELARRQRHPELSQ GEAVASVIYRTLAGLLPHLYDDKRSLRVPK RPINTPVVSISVHDDEGTAWLLOHTSGRAGG EVVFRNESHVSCQCHHMTSFALMDVSRRE GEBLAVSTYLAGLLPHLYDGKSLRVPK RPINTPVVSISVHDDEGTVULDFISLAGLUFLLGNQA DLFFACTVAILLHFLYLTFSWALLEAHLY RALTEVRDVNTGPMRPYYMLGWGVPAFTTG LAVGLDPEGGYGNPDFCWLSTYDLIWSFAGP VAFAVSMSVFLYLLAARSCCALQUFLLGNQA DLFFACTTAALLLLASVNSDTLL FHYLFATCNCQOPPELLSTVULLSVNSDTLL FHYLFATCNCQOPPELLSTVULLSSNSDTLL FHYLFATCNCQOPPELLSTVULLSSNSDTLL FHYLFATCNCQOPPELLSTVULLSSNSDTLL FHYLFATCNCQOPPELLSTVULLSSNSDTLL FHYLFATCNCQOPPELLSTVULLSSNSDTLL FHYLFATCNCQOPPELLSTVULLSSNSDTLL FHYLFATCNCQOPPELLSTVULLSFNXALL FHYLFATCNCQOPPELLSTVULLSFNXDSDTLL FHYLFATCNCQOPPELLSTVULLSFNXALL FHYLFATCNCQOPPELLSTVULLSFNXDSDTLL FHYLFATCNCQOPPELLSTVULLSFNXALL LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSCSPSYTPFLLR EESALNPGQGPGGGGGGGRCACCGGREKG PVSGLQPSFAVLLLLSATSRLALLSNSDTLL FESSALNPGQGPGGGGGGGGRCACCCGRFKCQ FESSALNPGQGPGGGGGGGGGGGCGCGGGGGRGRAFCACGGGGGGGGGGGGG							
YNGRPNEKIDFVALEVIQEQVQLTFSAGEST TTVSFFVPGGVSDGQWHTVQLKYYNKPLLG QTGLPQGPSQKVAVVTVDGCDTGVALRTGS VLGNYSCAA\QGTQGGSKKSLDLTGFLLLGG VPDLPESFVYRMGVGCMRNLQVDSRHDM ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAPSCECPLFGCKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQATTRGRSTITLQLREGHVMLSVEGTGL QASSLRLEPGRANDGDWHIAQLALGAIGGP GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV GGGPGPAGQVARGFGCQVRVSDTPGVN SLDPSHGGSINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCCDPGYGDNCTNVCDLNFP EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGPDPDC NKTSGGCHCKENHYRPGSPTCLLCDCYPTG SLSRVCDPEDQCPCKPGVTGRCCDRCDNPP AEVTTNGCEVNYDSCPRAEAGWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALILRINATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSFTTYTNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVPR ETTPVVNPAGPGBAGGPELARRQRRHPELSQ GEAVASVITYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVTRNESHYSCQCNHMTSFAVLMDVSRRE MGELPLKTLTTVALGLAGLAQLVFLLGINQA DLPPACTVIALLHET_LTLALLLIFFFLTLL RILRSNQHGRRNLTAALGLAQLVFLLGINQA DLPPACTVIALLHET_LTTSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSTYDTLIWSFAGP VAFAVSMVFLYLLAASLLALALLSVNSDTLL FHYLFATCNCQOPPHGLGRACLCAGREFKAQ PVSGLOPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCQOPPGLGRAKLCAGRGFEKKG PVSGLOPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCQOPPGLGRAKLCAGRGFEKKG PVSGLOPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCQOPPGLGFRKLCTGGRFKAQ RLYQPYGDSAOSLHSTSRSGKSQRSYFTPFLLR							
TITVSPFVPGGVSDGQWHTVQLKYYNKPLLG QTGLPQGPSEQKVAVVTVDGCDTGVALRGS VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG VPDLPESFPVRWRQFVGCMRNLQVDSRHDM ADFIANNGTVPGCFQAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQAITRGRSTITLQLREGHVMLSVECTGL QASSLRLEPGRANDGDWHHAQLALGAIGGP GHAILSPDYQQGAEGNLGPRLHGHLSNITV GGGPCPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGFYCET RDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSFTCLLCDCYPTG SLSRVCDPEDQCPCKPGVIGRQCDRCDNFF AEVTTNGCEVNYDSCPRALEAGIWWPETRFG LPAAAPCPKGSFGTAVRHCDEHGWLPPNLF NCTSITTSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHTTENLLRVGSALL DIANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPTTVTENVISVVRLDKGN FAGAKLJRYEALRGCQPPDLETTVILPESVPR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIVXTLAGLLPHNYDPDKRSLRVPK RPINTIPVVSISVHDGELLPRALDKRVTVOFFR LLETEERTKPICYFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRE NGELIPLKTLTYVALGVTLAALLLTFFFTLTL RILRSNOHGRRRNLTAALGLQLVFLLGINQA DLPFACTVIALLHFLYLCTFSWALLDALHLY RALTEVRDVNTCPMRPYYMLGWCVPAATTG LAVGLDPEGYGNPFCWLSIYDTLIWSFAGP VAFAVSMSVFLYTLAARASCAAQRQCFEKKG PVSGLQPSFAVILLLSATWLLALLSVNSDTILL FYLFATCNCQGPFTLSYVVLSKEVRAKK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQNYGDSAGSLHSTSRSOKSQPSYTFLLR							
QTGLPQGPSEQKVAVVTVDGCDTGVALRGS VLGNYSCAAVQGTQGGSKKSLDLTGPLLLGG VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQATTRGRSTITLQLREGHVMLSVEGTGL QASSILEPGRANDGDWHHAQLALGAIGGP GHAILSFDYQQRAEGALGPRLHGLHLSNITV GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC BHQSVCTRKPSAPHGYTCECPPNYLGPYCET RDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDCDNPF AEVITNGCEVNYDSCPRAEAGIWWPRITRG LPAAAPCPKGSFGTAVRHCDEHRGWVPRITRG LPAAAPCPKGSFGTAVRHCDEHRGWVPRITRG AFWTTNGCEVNYDSCPRAEAGIWWPRITRG LPAAAPCPKGSFGTAVRHCDEHRGWVPRITRG AFWTTNGCEVNYDSCPRAEAGIWWPRITRG LPAAAPCPKGSFGTAVRHCDEHRGWPNLFN NCTSITFSELKGFAERLQRIPESGLDSGRSQQL ALLLRNATQHTAGYFGSBVVAYQLATRLL AHESTORGFGLSATQDVHFTENLIRVOSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTINISVVRLIDKGN FGAKLLPRVEALRGCOPPDLETTVILIPESVPR ETPPVVRPAGPGEAQEPEELARQRRHPELSQ GEAVASVINTRLAGCOPPDLETTVILIPESVPR ETPPVVRPAGPGEAQEPEELARQRRHPELSQ GEAVASVINTRLAGLQLYPLLGDWDSKRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILETERKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILETSRQHGRRNLTAALLLTFFFLTLL RILETSRDWTGPRRFYYMLGWGYPAFTTG LAVGLDPEGYGNPFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVILLLLSATWLLALLSVNSDTLL FHYLFATCNCQGPFICKVISTYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVILLLSATWLLALLSVNSDTLL FHYLFATCNCQGPFICKYCTLSYVLISEVSWCRALK LACSKKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSOKSQPSYTFLR	1 1						
VLGNYSCAAVGGTQGGSKKSLDLTGPLLLGG VPDLPSFPVRMRQFVGCMRNLQVDSRHIDM ADFIANNGTVPGCCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQAITRGRSTITLQLREGHVMLSVEGTGL QASSLRLEPGRANDGDWHHAQLALGAIGGP GHAILSPDYQQGAEGNLGPRLHGHLSNITV GGDPPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGGSINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHTYRPGSFTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDCDNFF AFVTINGCEVNIDSCPRAEAGIWWPRITFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITTSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGFSDVXVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELQQTEGGTAWLQHYEAYAS ALAQNMRHTYLSPFTIVTPNVISVYLDLKGN FAGAKLPRVEALRGEQPPDLETTVLLPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIVRTLAGLLPHNYDPDKRSLRVPK RPINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSQCQNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLITFFFTLL RILRSNOHGRRNLTAALGLAQLVPLJCBNOA DLPFACTVIAILHFLYLCTFSWALLEAHLLY RALTEVRDVNTGPMRPTYMLGWGVPAFTTG LAVGLDPEGYGNPFOFCWLSTYDTLINSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVILLLLSATWLLALLSVNSDTLL FHYLFATONCQGPFGLGGIFGRALCGPSPVPFLIR EESALNPGNQGGBRACCSQPSYPFFLIR EESALNPGNQGGBRACCSQPSYPFILR	1 1						
VPDLPESFYRMROFVGCMRNLQVDSRIDIM ADPIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQAITGRSTITLQLREGHVMLSVEGTGL QASSLRLEPGRANDGDWHHAQLALGAIGGP GHALLSPTYQQRAEGNLOPRLHGLHLSNITV GGPGPAGGVARGFRGCLQGVRYSDTFEGVN SLDPSHGESINVEQGCSLPPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC BHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGPDPDC NKYSGECHCKEAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGPDPDC NKYSGECHCKEAPHGYTCECPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGPDPDC NKYSGECHCKEAPHGYPRGSPTCLLCDCYPTG SLSRVCDPEDQQCPCKPGVIGRQCDRCDNPF AEVTTNGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELQQTEGGTAWLLQHVEAYAS ALAQNMRHTYLSFFTIVTPNIVSVYRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPELARRQRRHPELSQ GEAVASVITRTLAGLLPHNYPDKRSLRVFK RPINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSLVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIALLHFLYLCTTSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFTTG LAVGLDPEGGYGNPDFVYYMLGWGVPAFTTG LAVGLDPEGGYGNPDFVYYMLGWGVPAFTTG LAVGLDPEGGYGNPDFCVLSTVDTLIWSFAGG VAFAVSMSVFLYILAARASCAAQRGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLTATCNCIQPFFLSYVVLSKEVRKAK LACSRXPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLISTSRSGKSQPSYFFLIR	i I						VI CHYSCA ALCOTOCCSYVELD TOST LLCC
ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGISLPISQPWYLSLMFRTRQAD GVLLQAITRGRSTITLQLEGHVMLSVEGTGL QASSIRLEPGRANDGDWHHAQLALGAIGGP GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV GGPGPAGGVARGFRGCLQGVRVSDITEGGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTTCCPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGPDPDC NKTSGECHCKENHYRPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKFOVIGRQCDRCDNFF AEVITNGCEVNYDSCPRAEAGIWWPRITRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTORGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSFTIVTPNVISVVRLDKGN FAGAKLPRYEALRGEQPDLETTVLLESVFR ETPPVVRPAGPGAQEPELARRQRRHPELSQ GEAVASVIYRTLAGLLPHNYDDFKRSLRVFK RPINTPVVSISVHDDEELLPRALDKPVTVOFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGGVTLAALLLTFFFTLLL RILRSNQHGFRNLTAALGLAQVFLIGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGYAFITG LAVGLDPEGGYGNPDFCWLSIVDTLIWSFAGP VAFAVSNSVFLYILAARASCAAQRGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTIL HYYLFATCNCIQOPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPY ADG RLYQPYYGDAGGLGGFGRKGCFLGRFKNDQQ				, ,			VPDI PESEDVINAD OEVICOMDNI OVIDODITIDIA
VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSIPISQPWYLSLMFRIRQAD GVLLQAITRORSTITIQLREGHVMLSVEGTGL QASSLRLEPGRANDGDWHHAQLALGAIGGP GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV GGPGPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCCKPGVIGRQCDRCDNFF AEVTTNGGEVNYDSCPRAEAGGWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHFELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLIGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFITG LAVGLDPEGVGNPDFCWLSIYDTLIWSFAGP VAFAVSNSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL HTYLFATCNCIQOPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSYYADG RLYQPYYGDSAGSLLSTSRSGKSQPSYPFLLR	i i						
LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQAITBGRANDGDWHLAQLALGAIGGP GHAILSPDYGQQRAEGNLGPRLHGLHLSNITV GGIPGPAGGVARGFRGLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNFF AEVITNGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVHCDEHRGWLPPNLF NCTSITFSELKGFAERLQNNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELQOTTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGFGAQEPEELARRQRHFELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPINTFVVSISVHDDEELLPRALDRFVTVQFR LLETERTKPICVFWNHSILVSGTIGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPKTLTYVAGLGVTLAGLAUTLLIFFFITLL RILRSNQHGRRNLTAALGLAQLVFLLGINQA DLFFACTVIALLHFLYLCTFSWALLEALHLY RALTEVRDVNTQPMRYYMLGWGVPAFTIG LAYGLDPEGYGNPDFCVLSIYDTLIWSFAGP VAFAVSMSVFLYLAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSLVNSDTIL HYLFATCNCLQGPFIFLSYVVLSKEVKRALK LACSRKPSPDPALTTKSILTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPGJQGPGLGGIPGRLCFLGRFKDQQ	1 1						
GVLLQAITRGRSTITLQLREGHYMLSVEGTGL QASSLRLEPGRANDGDWHHAQLALGAIGGP GHAILSPDYGQRAEGNLGPRLHGLHLSNITV GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGCPNCDVSKGFDPDC NKTSGECHCKENHYRPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVTINGCEVNYDSCPRAIEAGIWWPRITRG LPAAAPCPKGSGGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRTYLSPTTYTPNIVISVVRLDKGN FAGAKLPRYEALRGCQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVITYRTLAGLLPHNYDPDKRSLRVFK RPIINTPVVSIDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSQCCNHMTSFAVLMDVSRRE NGELPLKTLTYVALGVTLAALLLTFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLFFACTVIAILHFLYLCTFSWALLEALHLY RALTEVRDVHTOPMRPYYMLGWGVPAFTTG LAYGLDPEGYGNPDFCWLSTYDTLIWSFAGP VAFAVSMSVFLYTLAARASCAAQRQGFEKKG PVSGLQPSFAVLLLAARASCAAQRQFEKKG PVSGLQPSFAVLLLAARASCAAQRQFEKKG PVSGLQPSFAVLLLAARASCAAQRQFEKKG PVSGLQPSFAVLLLAARASCAAQRQFEKKG PVSGLQPSFAVLLLAARASCAAGRQOFFEKKG PVSGLQPSFAVLLLAARASCAAGRQOFFEKKG PVSGLQPSFAVLLLAARASCAAGRQOFFEKKG PVSGLQPSFAVLLLAARASCAAGRQOFFEKKG PVSGLQPSFAVLLLAARASCAAGRQOFFEKKG PVSGLQPSFAVLLLAARASCAAGRQOFFEKKG PVSGLQPSFAVLLLAARASCAAGRQOFFEKKG PVSGLQPSFAVLLLASTWLLALLSVNSDTLL FHYLFATCNCQPFILSTVVLSKEVKKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLR EESALNPGLQGPGLGGPGRLCGTFKDQQ							
QASSLRLEPGRANDGDWHHAQLALGAIGGP GHAILSFDYQQRAEGNLGPRLHGLHLSNITV GGPGPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSFTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL AILLRNATQHTAGYFGSDVKVAYQLATRLL AHESTORGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRTYLSPFTLYTPNIVISVVLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHFELSQ GEAVASVIYRTLAGLLPHNYDPKRSLRVPK RPINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGELPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILHFLYLCTFSWALLEAHLY RALTEVRDVNTGPMRPYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYLLARASCAAQRQGFEKKG PVSGLOPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQQPFIFLSYVVLSKEVRKALK LACSRKPSPPPALITIKSTLTSSYNCPSFYADG RLYQPYYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNFQQGPGLGGIFGR/LCFLGRFKDQQ	1 1				ľ		
GHAILSFDYQQQRAEGNLGPRHGLHLSNITV GGIPGAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDCNDNFF AEVTTNGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLIRNATGHTAGYFGSDVKVAYQLATRLL AHESTQRGGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTTVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPELARRQRRIFELSQ GEAVASVIIYRTLAGLLPHNYDDPKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGGELPLKITLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLFFACTVIALLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILLAARASCAAQRQGFEKKG PVSGLOPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQOPFFLSYVVLSKEVRKALK LACSRKPSPDPALITKSTLTSSYNCPSPYADG RLYQPYYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNFQQGFELKGDQQFFICKGDQQ							
GGPGPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVITNGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITTSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGCQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPELARRQRRHPELSQ GEAVASVITYRTLAGLLPRNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGRRNLTAALGLAQLVFLLIGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RAALTEVRDVNTGPMRPYYMLGWGVPAFTIG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQQPFIFLSYVVLSKEVRKALK LACSRKPSPDPAALTIKSTLTSSYNCPSFYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLIR EESSALNPGQGPPGLGGPGRLCFLGRFKDQQ							
SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTKKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDNPF AEVITINGCEVNYDSCPRAEAGIWWPRITRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPTTIVTPINVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPELARRQRRHPELSQ GEAVASVITYRTLAGLLPHNYDPDKRSLRVPK RPINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFI-TILL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTTSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFITG LAVGLDPEGYGRPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPAALTIKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPGQGPPGLGGPGRALCFLGRRKDQQ]		ļ	j			
CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RDQPCPRGWWGHPTCGPCNCDVSKGPDPDC NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVTTNGCEVNYDSCPRAÆAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPINTPVVSISVHDDEELLPRALDKPVTVOFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLFFACTVIALLHELYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFACTVIALLHELYLCTTSWALLEALHLY FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTRSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLIR EESALNPG\QOPGLGGIGRRLCFLGRFKDQQ	1		i				
EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNFF AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTTVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPINTTVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNHOFGRRNITAALGLAQLVFLLGINQA DLPFACTVIALLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSTYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTILL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALITIKSTITTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLIR EESALNPG\QOPPGLGGIPGR/LCFLGRKKDQQ	ļ]					
RIDOPCPRGWWGHPTCGPCNCDVSKGFDPDC NKTSGECHCKENHYRPGSFTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNFF AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITTSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTORGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGELPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFILGINQA DILPFACTVIAILLHFLYLCTFSWALLEALHLY RAALTEVRDVNTGPMRFYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYPFPLR EESALNPG\QGPPGLGGIPGR/LCTLGRFKDQQ	1 1						
NKTSGECHCKENHYRPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVTINGCEVNYDSCPRAEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RAALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQQPFILSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCTLGRFKDQQ	1	ľ		Ì		ł	
SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVTTNGCEVNYDSCPRAEAGIWWPRITRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPGNQGPPGLGGIPGR/LCFLGRFKDQQ				· i			
AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPTITVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGYPAFTTG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QQPPGLGGIPGR/LCFLGRFKDQQ							
LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR RPIINTPVVSISVHDDEELLPRALDKPVTVQFR RPIENTFVVSISVHDDESLLPRALDKPVTVQFR RPIENTFVVSISVHDDESLLPRALDKPVTVQFR RPIENTFVVSISVHDDESLLPRALDKPVTVQFR RPIENTFVVSISVHDDESLLPRALDKPVTVQFR RPIENTFVVSISVHDDESLLPRALDKPVTVQFR RPIENTFVVSISVHDDESLLPRALDKPVTVQFR LIETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QQPPGLGGIPGR/LCFLGRFKDQQ]	j	j	ľ			
NCTSITFSELKGFAERLQRNESGLDSGRSQQL ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYYGDSAGSLHSTSRSGKSQPSYIPFILR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	[[ſ	1		
ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFTTG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ			ŀ	ļ	ļ		
AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFFILSYVVLSKEVRKALK LACSRKPSPDPALITKSILTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYPFFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	.	İ	ļ	}			
DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ							
ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRPYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYPPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ		J	j	j	J		
FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ		ļ			1		
ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ		ŀ	ł		l		FAGAKLPRYEALRGEOPPDLETTVILPESVFR
GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLALLSVNSDTLL FHYLFATCNCIQGPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQPYGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ		-	1			1	
RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ		ł	ł	}	1		GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ				-	1	İ	
EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ]	1	1	ļ	
NGEILPLKTLTYVALGVTLAALLLTFFFLTLL RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ			1	1	l	1	
DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYPPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	1		- 1	ĺ	1	1	
DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYPPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ			- 1	1	l		RILRSNQHGIRRNLTAALGLAQLVFLLGINOA
RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP/YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ					1	İ	
VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ			- 1		. 1	1	
PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ				1		ŀ	LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	-		Į	İ		1	
FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	1	1	1	1		1	PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL
LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	ĺ	ĺ	1	1		1	
RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFILR EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	. 1	ļ	i	j		ŀ	
EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ	'	ł	İ	ļ		.	
		ļ				•	
		ļ	1		·		H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EEEEEEEEAAFPGEQGWDSLLGPGAERLPLHS TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP EERLRENGDALSREGSLGPLPGSSAOPHKGIL KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE
425	1775	A	3429	155	1417	GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA GTVDEDSSGSEFLFFNFLH GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSS RAASVREAEDAPLQPASIHPVSQGSRGPEGSL
						GSAECLPGDPLGARRATRAHSPVPGPPPSLPA AGTAVKRGLQPG*GA/GATSTPGTGAATGGL CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV* KREFQRGPWAGMVILHRISAADPARAPGPDS NLQSALQQPATGCSEPAAVYSPPIGLWGA**P EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI YELLENGQRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLFCRTLEDILADAPESQN NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS
426		A	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTG\SHDFAINFDPENPECEGK RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HPRPVFI\EISGVIASYRRCLPQIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP
427	1777	A	3446	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPPQAQPLLPQPQPPPPPPPPPP

SEQ ID No. of nucl- ording peptide solide in uncloodide peptide solide in uncloodide peptide solide in uncloodide per per per per per per per per per pe							
peptide contide sequence uence uence uence uence peptide corresponding to lost at mino acid residue of peptide sequence peptide sequence peptide sequence sequence sequence peptide sequence peptide sequence sequence peptide sequence sequence peptide sequence sequence peptide sequence sequence peptide sequence sequence peptide sequence peptide sequence sequence peptide sequence peptide sequence peptide sequence peptide sequence sequence pertia sequence							
sequence USSN plays corresponding to first amino acid residue of peptide sequence Polyano corresponding to first amino acid residue of peptide sequence Polyano corresponding to first amino acid residue of peptide sequence Polyano corresponding to first amino acid residue of peptide sequence Polyano corresponding to first amino acid residue of peptide sequence Polyano corresponding to the sequence of peptide sequence Polyano corresponding to the sequence of peptide sequence of peptide sequence Polyano corresponding to the sequence of peptide sequence of			hod .		, , ,		,
uence 1944	ոսշվ-	peptide	ļ l		nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
uence 1944	eotide	seq-		USSN	location	corresponding	
Service Serv	seq-				correspondi		
mino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per control of the peptide seque							
residue of peptide sequence Y=Tyrosine, X=Unknown, =Sipp nodon, y=possible sequence Tespegoty Spinny Deligopoy (Notation of the peptide sequence)	1		j				
peptide equence Tieypeegy/SDILNYDHORPQVRGATALLC. CTILCSLSSRSPHYGDWMGTHRITLOTHESL. ADCPPLIRKTIKDESSYTCKLACTAVRNCVM. SLCSSYSPELGULIDUVILINSSYVLVRTEL LETLAEDFRIVSTLEAKARMLHRGAHHYTGL. LELQERVLNNVVEILLDEDPRYWHVAAASL. RILVPKLFYKCDQGADPVVAVARDQSSVYL. KLLMBETQPPSHEVSTTIRTRGAVHAAASL. RILVPKLFYKCDQGADPVVAVARDQSSVYL. KLLMBETQPPSHEVSTTIRTRGAVHAAASL. RILVPKLFYKCDQGADPVVAVARDQSSVYL. KLLMBETQPPSHEVSTTIRTRGAVALASTIRTGCECE. ALCALSTAFPVCIWBLGWHCOVPPLSADSER. KSCTVGMATMITLLSSAWPPDLSAHDAVA. GEEVWPALGDRALASPWKPQLFSHLLKVINC. ANVIDDVAPPAIKAALPSTITPSLSTRPKL. GEEVPPALGDRALASPWKPQLFSHLLKVINC. ANVIDDVAPPAIKAALPSTITPSLSTRPSLSPRRK. GEEKPPGQASVPLSSKKSGSAAASRQSDTS. GOVTTSKSSSLGSFYHLFSYLKLHDVKATHAA. NYKYLDLONSTEREFGOFLRSALDVSQLELL. ATLQDIGKCVEELGYLKSGTSRPMMATVC. VQULKTLRGTHALSGYGFLSSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYGFLSAMSTGSGRA. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYTCDDASCR. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYTCDDASCR. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYTCDDASCR. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYTCDDASCR. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYTCDDASCR. QRLGSSSVBPGLHBYCGMATVC. VQULKTLRGTHALSGYTCDDASCR. QRLGSSSVBPGLHBYCGMATVC. VGLKCMATVC. VQULKTLRGTHALSGYTCDDASCR. QRLGSSPFTLDAVGCMATCHALBYCG							
miclotide insertion TTEYPEROY VSIDL NYDHGDPQVRGATALLC GTILCSILSRSRPHODWAGTRIT.TGNTFSI. ACCIPILIRIT.LKGESSVCKLACTAVRICVM SILCSSSYSELGLQLIDVIT.RRNSYWLVRTELL LETLAGEDFRI.VSTLEAKAENLIRGAGHPYTOL LEKLGERVI.NVVIHLLGDEDPRVEHVAAASL RILVPKLEVCQQOQADPVVAVARDQSSVYL KILLMEDRIVSSISSWSTITRIYRGYNLLPSITD VMENNLSRVIAAVSBLITSTTRALTIGCCE ALCLISTAPPVCIWSI.GWRICGYPPLSASDESR KSCTVGMATMIT.ILSSAWPFUDLSABDESR KSCTVGMATMIT.ILSSAWPFUDLSAAHQDAL LLAGNILAASAPKSIRSSWASEEANPAATK QEEVWPALGDRALVPWOQLIFSHLIKVNIC AHVLDDVAPPGAIKAALPSLTNPPSLSPIRRK GEKEPPGGASVPILSKYGSEASASRQSDTS GPVTTSKSSSI.GSFYHLPSYLKI.HDVLKATHAA NYKVTLDLONSTEKEFGGFRSAALDVIS.QUEL ATILQDIGKCVEELIGVILSGGFRSAMMATVC VQULKTLFOTHALSOPTOLSOPELS SUSPICITY OF THE STANDARD AND AND AND AND AND AND AND AND AND AN	1		l i	[Soderen	
TTEYPEROYN SDILNYDHORPOROVATABLE GTILCELL SASKPHY DOWNOTHER LICENTS: ADCPLIRKTILKDESSVTCKLACTAVRNCVM SLCSSSYSELGLQLIDVILTURISSYVLVATEL LETLAEDFRIVST EAKAEMLHRGAHEYTGL LELQERVINNVVHILLDEDPRVHAVAASL RLVPKLFYKCDQQADFVVAVARDQSSVYL KLLMEHCTOPSHEVSTTIRTRGVALLSTID VTMENNLSRVILAVSHELITSTRALTFGCCE ALCLESTAPPCVINSLG WHICKOPPLASADESR KSCTVGMATMILTLSSAWPEDLISAHDDAL LAGNALASAPKISLG WHICKOPPLASADESR KSCTVGMATMILTLSSAWPEDLISAHDDAL LAGNALASAPKISLSSWASEEGAPPAATK QEEVWPALGDRALVPMVSQLFSIRLKVNNC ANVLDDVAPOPARKALPELTOPPSISPIRK GKEKEPGBQASVPLSPKKOSEASAARQSDTS GPVTTSKSSLGSFYRLFSYLKLHVLKATHA NYXYTSLSNSSLGSFYRLFSYLKLHVLKATHA NYXYTSLSNSSLGSFYRLFSYLKHDVLKATHA NYXYTSLSNSSLGSFYRLFSYLKHDVLKATHA NYXYTSLSNSSLGSFYRLFSYLKHDVLKATHA NYXYTSLSNSSLGSFYRLFSYLKHDVLKATHA NYXYTSLSNSSLGSFYRLFSYLKHDVLKATHA NYXYTSLSNSSLGSFYRLFSYLKHDVLKATHA NYXYTSLSNSSLGSFYRLFSYLKNSSAGAA QRLGSSSYMPLSYLKOSTOLKY GVOLKTLFGTNASOFDCISSAPSKAGAA QRLGSSSYMPLSYNKOSTOLKY NITTYTORADONATHTHTQALADA SLRMMVQAEQBNDTSWOPTUKKYSTQLKY NITTYTORADONATHTHTQALADA SLRMMVQAEQBNDTSWOPTUKKYSTQLKY NITTYTORADONATHTHTQALADA VXYTSLRYFRYSKGIGHKRIDLCDGMASGR KAVTHAPALQPVHEILFRYCTNREADGMASGR KAVTHAPALQPVHEILFRYTNREADGMASGR KAVTHAPALQPVHEILFRYNTRIADAGKE LETQKEFVYSMLLIKLGYFSYLCKUCHHINKS QCHERSENGKNLPETSRFLLQUVGLLEDIVT KRIGHESTSPYLLSCTVNREADGDSTSTILE BHSSGKQNKNLPETSRFLLQUVGLLEDIVT KOLKVEMSQQHITPYNRW ARVQQTTRABILSSTRTUNGLTHTYNRW ARSWGTTHAASTSTVQLWISGILALRVLSQSTED VISRIGELSFSPYLLSCTVNREADGDSTSTILE BHSSGKQNKNLPETSRFLLQVGLLICHTYNRW ARVQQTTRABILSSTRTUNGLHINKS QMFRRTAAATRIFSDCGGGSFYTLDSIALR ARAMITHPALVLL WCCHLLIVNHTYNRW ARVQQTTRABILSSTRTUNGLHINGS UCKRESSQCHTNSCSSTPTMLKKTLQCLICH HLSQSGAVLTLYVDRLLCDFRYLARMVDIL ARRYBEMLSAHLGSLSTFYMLKTLQCLICH HLSQSGAVLTLYVDRLLCTFRYNLARMVDIL ARRYBEMLSLAPCLSCKMSSGGGGREGISCLA ARLGMCKNEUNGRALLFCDYVCQHLHISKS QCWTRSDSALLEGAELVNRTAEDMNAFM MISEPILSLAPCLSCKMSSGGGGREGISCLA AREUTLARVSGTVQLPAVHHYOPELPAPE ANYSKINDLEGGAAUSKLPLAALAQY UVVSKLPSHLHLPPEKENDVKFVVATLEAL SWHLEBEGGREGORDTAAFE WSVSSTEPVHAGSLLVKVFRILTWYNLOUNG UVSVSSTEPVHAGSLLVKVFRILTWYNLOUNG QESPPEEDTERQNNLAVAGLOCCLALQCHA WSVSSTEPVHAGSLVKVKFVATLA	1		ļ i		1 :		• •
GTILICSIL SERSPHYGDWAGTRIT.TIGNTES. ACCIPILLRITLKDESSYCKIL.ACTAYRICVM SICSSSYSELGLQLIDVITLRINSSYWLVRTEL LETLA ELIDERIL VSTIL EAKARDH.HRGARHYTGL LIKLQERVLNNVVIHLLQDEDPRVAHVAAASL RILVPKLIPKCDQQ.ADPVVAVARDDQSSVYL KILLMEFTOPPSIFSVSTITRIYRGVNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLISTAFPVCIWSLGWEIGCVPPLASADESR KSCTVGNATMITLISLSAWPPULSAAPDGAL LIAGNLLAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPMVSQLFSHLLKVINIG ARVLDDVAFGPIARALPSILTNPSSFIRKK GKEREPGEQASVPLSYKKGSEASAASRQSDTS GVYTTSKSSSLGSFHALPSYNLLEHDVIKATHA NYKVTLDLQNSTEKFGGFI.RSALDVLSQUELL ATLQDIGCVEELIGVILSCFSERPMATVC VQQLLKTLEGTNLASGPGGLSSNPSKSQGA QRLGSSSVPGCLYNYCHSCFSERPMATVC VQQLKTLEGTNLASGPDGLSSNPSKSQGA QRLGSSSVPGCLYNYCHSCFSERPMATVC VQQLKTLEGTNLASGPDGLSSNPSKSQGA QRLGSSSVPGCLYNYCHSCFSERPMATVC VQQLKTLEGTNLASGPDGLSSNPSKSQGA QRLGSSSVPGCLYNYCHSCFSERPMATVC VQQLKTLEGTNLASGPDGLSSNPSKSQGA QRLGSSSVPGCLYNYCHSCFSERPMATVC VQQLKTLEGTNLASGPDGLSSNPSKSQGA QRLGSSSVPGCLYNYCHSCFSERPMATVC VQQLKTLEGTNLASGPDGLSSNPSKSQGA QRLGSSVPGCLYNYCHSCFSERPMATVC VQQLKTLEGGRIPHGVLQKVSTQLKT INTSVTORADKNAHHHRILFFELYKALKQ YTTTTCVQLQKQVLDLLAQLVQLKVNYCLL DSDQVTFGVNLGGFEVGQFTCSSALINNFF FLVLLSYERYYISKQIGGFKQICDCGMASGR KAVTHAPALQPYHUDLVLQTGNTAAGKE LETQKEVVVSMLLRLQYHQVLEWFILVLQQ QCKENEPKWRKLSRGQIDHIPVLAGGNAGKE LETQKEVVVSMLLRLQYHQVLEWFILVLQQ QCKENEPKWRKLSRGQLDILIPVLAGGNAGKE LETGKEVVVSMLLRLQYHQVLEWFILVLQQ QCKENEPKWRKLSRGQLDILIPVLRGSTED VIJSRIQGLSFSPYLISCTVINRLRQGDSTSTLE BHSGGGKINLPEETFSSFLLQLVGGILLEDIVY KQLKVEMSEQQHTFVCQGLGTLLMCLHIFKS GMFRRITAAATTLFSSFLLQLVGGLLEDIVY KQLKVEMSEQGHTFVCQGLGTLLMCLHIFKS QGVTTSRASSTRAATTLFSSFLLQLVGGLLEDIVY KQLKVEMSEQGGSGSTLLSGSTEDHALGHAR ARSMITHPALVLLWCQLLLLVNHTDYRWW ARVQGTFKRHSLSSTFALLGPWGGGSESPSILA ARLGACNRETVERGALLFCPFVCALAMDOL ACRRVEMLLAANLQSSMAQLPMELNROPY LQSGGLAGHQALVSLUPERFSLTMODSLSRS PVYSSIPLDGGGGNSTLLTFFRTAKATLFSPX ASGLFGAQGACHALVSLUPPREDNAMM MSSEPPLSLLAPCLSGMSVFWFWATALAQL WSCHARTIMATAGLAGGACHALTGRAAAQPOL CORRESPECTORNYLAVQGLGCCCLAGLGRINGF TAACEMYABCACAGACHALTGALAQY LVVVSKLSHLHLPPSKRDIVKFVVATLALA WHILEGEGRIDLQGAMVSKEBNATHHUNQAWD PYSLSSPATTGALLSEKELLLQNPTEGLGSMS	——	ļ			sequence		
ADCIPILIRATILKDESSVICKLACTAVRNCYM SICSSSYSELGIQLIDVITLRINSSYVLYRTEL LETLAEIDFRI VSTLEAKARNLHRGAHHYTGL LKLQERVINNOVIHLL DEDEPVEHVAAASI RLVPKLFYKCDQGADPVVAVARDQSSVYL KLLMEIDFPSHISTVSTITTRYGVILLSTID VTMENNLSRVAAVSHELITSTTRALTFGCE ALCLISTAFPVCWSLQWECGVPLSASDESR KSCTVQBATMILTILSSAWFHILDSAARQAL LAGNILAASAPKSIRSSWASEEANPAATK QEEVWPALGRALVPWEQLEFSHLKVNIC AHVLDDVAPGAKAALPSLTNPFSLSPIRKK GEKKEPDEQASVPLSPKGSEASAASRGSDTS GPVTTSKSSSLGSFYHLSVILKHDVLKATHA NYKVTLDLQNSTEKFGFURSAADVSQUEL ATLQDIGKCVEELGYLKSCFSREPMMATVC VQULKTLFGTNLASOPDELSSPRSSKOGRA QRLGSSSVRPGLYHYCPMAPYTHFTQALADA SLRIMVQAEOENDTSGWPDU,GXVTQLKT NLTSYTKNRADRNAIHHRILFPLVIKALKQ YTTTCVOLQKVQLDLLAQLVQLRVNCLL DSDQVFIGFVLKQFEVEVGQYRESEAIPMIFF FVLISYREYHSKGIGFKGKJOLOVSTQLKT NLTSYTKNRADRNAIHHRILFPLVIKALKQ YTTTTCVOLQKVQLDLLAQLVQLRVNCLL DSDQVFIGFVLKQFEVEVGQYRESEAIPMIFF FVLYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKYSKY	i			1	Į i		1
SLCSSYSELGLQLIDVLTINSSYWLVATE LETLAEDRIA/STLEAKAENLHRGAHHYTGL LKLQERVLINNVHILLGDEDPRVRHVAASL RLVPKLFYKCOGQADPVVAVADQSSYYL KLLMHETQPSHESVSTITRIYRGYNLLPSITD VTMENNLSRVLAN SHELITSTTALTGCCE ALCLISTAFPUCIWSLGWHCGVPPLSASDDSA KSCTVOMATMILTILSSAWPDLSAHODAL LAGNILAASAPKSLRSSWASEEANPAATL QEEVWPALODRAL VPMWGQLFSHLK VNIC AHVLDDVAPGPAIKAALPSLTNPFSLSPIERK GKEKEPGGQASVPLSPKKGGEASAASRQSDTS GVPTTSKSSSLGSTYHLPSYLKLBHOVLKATHA NYKVTLDLQNSTEKFGGFLSAASRQSTST GVPTTSKSSSLGSTYHLPSYLKLBHOVLKATHA NYKVTLDLQNSTEKFGGFLSAASRQSTST GVPTTSKSSSLGSTYHLPSYLKLBHOVLKATHA NYKVTLDLQNSTEKFGGFLSAASRQSTC ATLQDIGKCVEELGYLKSCSFSRPMAATVC VQQLLKTLFGTNLASQFDGLSYNSSSSGAA QRLGSSSVRGELYHCSWALTHITQALADA SLRNNVQAGQENDTSGWFDVLQKVSTQLKT NLTSYTNRNARAKNALHNIBLIFELPYKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLRGFYEVGGGFESALIPNIFF FLVLLSYERYHSKQDGPKRQLCDGHAASGR KAVTHAPALQPVDLLFVLRGTNAAGKE LETQKEVVVSMLLRLIQVHQVLEMFILVLQQ CHKERDEDWKRLSQAIDALRAQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLSSMF VTPNTMASVSTVQLWSGLALBLANGQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLSSMF VTPNTMASVSTVQLWSGLALBLANGSSFED IVLSRQELSFSFYLISCTVINRLRDGDSTSTL BISBGKQKNLPEETFSRFLLQLVGILLEDITLY KQLKVEMSEQQHTTYCOGLGTLLMCLHIFKS GMFRRTAAATRIFSBOCGGSFLUSICISHAR ARSMITHPALVLLWCGLLLLVHICHDSHLR ARSMITHPALVLLWCGLLLLVHICHDSHLR ARSMITHPALVLLWCGLLLLVHICHDSHLR ARSMITHPALVLLWCGLLLLNSKALDSHLP JERGGAVLLTVPRULLLTSHPWLOFLLSMRL ARSMITHPALVLLWCGLLLVHTRYRW ACQOPTRENISLLSHEPPYCOPLLSSHLP AAKLGMCNREIVRRGALLFCDYYCQNLLIDSH ARGGFIQAOSRCENLSTPMLKKTLQCLEGE LLSGGAVLLTVPRULLTSHPWLARMVDIL ACREVEMLAANLQSSMAQLPREEINCLOPPS LQSGGAVLTAYDRLLLAFCHYOPLLARALAQP LLVYSKLPSHLAIDGAALYQSLTHARALAQP AVWSLDBLFGGAALYGSLTPTLARALAQP LLVYSKLPSHLAIPPERGESTORALAQP AVWSLDBLFGGAALYGSLTPTLARALAQP LLVYSKLPSHLAIPPERGESTORALAGENSTENATHHY YQUKUA GESPFEEDTERTQINVLAVQAITSLVLSAMT VPVAGNAVSCLEGCROPPRICHLAKRNSGVA FLITLLINNISLARLPLVNSTTRVPLYWKLG GESPFEEDTERTQINVLAVQAITSLVLSAMT VPVAGNAVSCLEGCROPPRICHLARRNSGVA LSIERGVPSCEIQAMVSKERNATHHY YQVRL	1		l i		ļ i	l i	
LETLAEIDFRLYSTLAKABHTOLISCH LKLQRVLNNVHLLGEDEPRVRVWAAASL IRLYPKLFYKCDQQADPVVAVARDQSSVYLL KLLMHETQPPSHESVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTTGCCE ALCLISTAFPCURSIGWHCGVYSTLASDESR KSCTVOMATMILTILSSAWPFLDLSAHDDAL ILAGNILAASAPKSLRSSWASEEEANPAATX QEEVWPALGDBALVPMVEQLFSHLKVINIC AHVLDDVAPQPAIKAALPSLTNPPSLSPIRKX GKEKEPGQASVSLSFKKGGSASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHAA NYKVTLDLQNSTEKFGGFLRSAKORSOTS GPVTTSKSSLGSFYHLPSYLKLHDVLKATHAA NYKVTLDLQNSTEKFGGFLRSAKORSOTS GRUTTSKSSLGSFYHLPSYLKLHDVLKATHAA NYKVTLDLQNSTEKFGGFLRSALVSQUSTAL ATLQDICKCVEELGYLKSCFSREPMMATVC VQLLKTLFGTNLASQFDGLSSFNGQFA QRLGSSSVRFGLYHCFMAPYTHFTQALADA SLRNMVQAGEDEDTSGWFDVLKQKSTGLKT NLTSYTKNRADKNAHNHIRLFEFLVKALKQ YTTTTCVQLKQKVDLLLAQLVQCKVSTGLKT NLTSYTKNRADKNAHNHIRLFEFLVKALKQ YTTTTCVQLKQKVDLLLAQLVQLEWTVCLL DSDQVFIGFVLKQFFYEVGQFRESEAINFIF FLVLLSYERYHSGKGIGFKRIQLOCIGIMASGR KAYTHAPALQPIVEDLFVLRGTNKADAGKE LETQKEVVSMLLALLIQYHQVLLEMFLVLQQ CHKENEDKWKRLSRQIDDILPMLAKQOMH DSHEALGVLATTLELAPSSLRPVDMLLSSMF VTPNTMASVSTVOLWISGILALLRYLISSTED VLSRIGGLSFSFYLLSCTVINRLDOGTSTILE EHSEGKQKNLPEETFSFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFIKS GMFRITAAATRLFSDCGGGSYLLDSINKR ARSMITHFAALVILWCQULLLVNHTDYKWW AEVQQTFKRISLSSTKLLSFQNGSEEDISIAA AKLGMCNREIVRRGALLHCDYVCQNLHDSS HTWILNNHHQDLISJSHEPVOPOFISAVERNS AASGLFQAJQSRCENLSTPTMLKKHLQCLGSISSS HTWILNNHHQDLISJSHEPVOPOFISAVERNS AASGLFQAJGSRCENLSTPTMLKKHLQCLGSISSS PVSSHFLLGDGHVSLETVSPDKDWYVHLV SQCWTRSDSALLBEGALNNRTRAEDMNAFM MNSEPNISLLAPCLSLGMSEISGGQKSALFE AYWSKLINDLFGDAALYQSLPTLARALOQU LVVVSKLPSHLHLPREKEDIVKVATLEAL SWHLHBOPPLSJLDLQGAGCCCLALQUERY AYWSKLINDLFGDAALYQSLPTLARALOQU LVVVSKLPSHLHLPREKEDIVKVATLEAL SWHLHBOPPLSJLDLQGAGCCCLALQUENCY ACREMPENTSHLAPCSLYCOLFVHPILVWALLEAL SWHLHBOPPLSJLDLQGAGCCCLALQUENCY TAGCMVASKELGEGEVFKEFTVALTICALITYCPLWE QEESPEEDTERTQINVLAVQALTSLVISAMT VPVAAGPAVSCLEOQOPRNETLALTRIPLYQAWD GEESPEEDTERTQINVLAVQALTSLVISAMT VPVAAGPAVSCLEOQOPRNETLALTRIPLYQAMD PVFSLSPATTGALBEELLLQUPFRELGSMS	1 .			i i	l l	!	
LKLQERVLNNVHLLGDEDPRVRHVAARDQSSVYL KLLMHETQPSHESVSTITRIYRCYNLLPSIT VITMENNLSRVLAAVSBELITSTTRALTFOCCE ALCLLSTAFPVCIWSLOWHCGVPPLSASDESR KSCTVGMATMILTLLSSAWPFLDLSAHQDAL LRONLAASARKSIRSSWASEELANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVNIN AHVLDDVAPGFAKAALPSILTPSISPIRKK GKEKPGGQASVPLSPKKGSEASAASRGSTSTS GPVTTSKSSSLGSTYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQULEL ATLQDIGKCVEELGYVLKSCTSRFDWAATVC VQQLLKTLFGTNLASQFDGLSSNFSKSQGA QRLGSSSVRFGLYHCSTKATHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKLK NLTSYTKNRADKAHANHHLLFEPLVKLAKQ VTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVTGFVLLQCFYTEVSQFRESEALPNIFF FLVLLSYERYHSKQIGDPKQULCOGMASGR KAVTHAPPALQPIVDLFVLRGTNKADAGKE LETOKEVVSMLALRLIQYHQVLEMTLVLQQ CHKENEDKWKLISRQIADDLPMLAKQQMH DSHAELGVLANTLELAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTELE EHSEGKQKKNLPEETFSRFLQLVGILLCIBVKS GMFRRTAAATRLFRSDGCGGSFYLDSIALR ARSMITHPALVLU WQGLLLVHNTDYRWW AEVQOTPKRISLSTKLLSPQMSGEEGEISHL RARMTHPALVLU WQGLLLVHNTDYRW AEVQOTPKRISLSTKLLSPQMSGEEGEISHLAR AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDVRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING AKAGMCNREIVRGALIFLFOUNTDRING ANDERSHLIARDGROUNDRING MNSEFNISLLACCIS.GMSEISGGQKSALFE AYWSKLINDLEGDAALVGSLPTLARALAQU UNVSKLPSHI-HIPPEKEDDIVKVVATLEAL SWEILHBQDFLSDLAGGLOCCLAQLQIGI WSVSCASDAALGREFETWATLGVU-VYQFLVME QEESPEEDTERTQINVLAVQAITSLVLSAMT TAACERVAEMVSELGSVLALDIKRRSVYYA TAACERVAEMVSELGSVLALDIKRRSVYYA TAACERVAEMVSELGSVLALDIKRRSVYYA TAACERVAEMVSELGSVLALDIKRRSVYYA TAACERVAEMVSELGSVLALDIKRRSVYYA TAACERVAEMVSELG	1		ļ i		ļ		L = = - · · · · · · · · · · · · · · · · ·
IRLÚPKLFYKCDQGQADPVAVARDQSSVYLLPSITID KILMHETQPSHESVSTTRIYRQYNLLPSITID VTMENNLSRVIAAVSHELIJSTTRALTFGCCE ALCLLSTAFPUCHSLGWHGCVYSYNLLPSITID VTMENNLSRVIAAVSHELIJSTTRALTFGCCE ALCLLSTAFPUCHSLGWHGCVYSADSSSR KSCTVOMATMILTILSSAWPFILDSAHDDAL ILAGMILAASSAPSKISGSWSBEDAPATK QEEVWPALGDRALVPMVEQLFSHLLKVNIC AHVLDDVAPGPAIKAALPSLTNPTSLSPIRK GREKEPGGQASVYLSPKKGGSASASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHAA NYKVTLDLQNSEKFGGFLRSALVSQUEL ATLQDICKCVEELIGYLKSCFSREPMMATVC VQLLLKTLFOTHLASQFPGLSSNFSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAGEDDTSGWFDVLKSGTGLK QVITTTTCVQLCKQVDLLAQLVQLSVTQLKA NLTSVTKNRADKNAHNHIRLFEPLVKALKQ YTTTTCVQLCKQVDLLAQLVQLDLAQLVQLSVTQLK NLTSVTKNRADKNAHNHIRLFEPLVKALKQ YTTTTCVQLCKQVDLLLAQLVQLDCHMASGG KAYTHAPALQPIVEDLFVLRGTNKADAGKE ETGEVVVSMLLRLIQYHQVLEMFULVLQ CHKENEDKWKRLSRQIADULPMLASWMF FLVLLSVERYHSKGIGPKRIQLCGIMASGG KAYTHAPALQPIVEDLFVLRGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFULVLQ CHKENEDKWKRLSRQIADULPMLASWMF VTPNTMASVSTVOLWISGILAILRVLISSTED VLSRGGLSFSFYLLSTVNRRLDGSTSTLE EHSEGKQKNLPEETFSFFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAATATLFSDGCGGSFYLDSINKR ARSMITHFALVILWCQULLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSFQMSGEEDSILA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWUNNHQDLISSSTKLLSFQMSGEEDSILA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWINNHQDLISSSFKLLSFQMSVPRINARM ARSGAVILLTYDBLLAFRINGDNAFN ARSGGAVILTYDBLLAFRINGDNAFN MNSEFNISLLAFCLSLGMSEISGGQKSALFEN AASGLFQAJQSRCENLSTFTMLKKTLQCLES PVSSHFLIGDGHVSLETVSPDKDWYHLVK SQCWTRSDSALLBEAGLINNFRAEDMNAFN MNSEFNISLLAFCLSLGMSEISGGGKSALFA AFLYLLARVSGTVOQLPAVHEYPFQFLPAFA AAYWSLLNDLGGDAALYQSLPTLARALDQI UNVSKLSBGAVILTYDPLKCHIVKVVATLEAL SWHLHEDPISLDLQAGLDCCLALQUPL WSVYSLSPHLHEDPISLDLQAGLDCCLALQUPL WSVYSLSPHLHEDPISLDLQAGLDCCLALQUPL WSVYSLSPHLHEDPISLDLQAGLDCCCLALQUPL WSVYSLSPHLHEDPISLDLQAGLDCCLALQUPL WSVYSLSPHLHEDPISLDLQAGLDCCCLALQUPL WSVYSLSPHLHEDPISLDLQAGLTCCLALQUPL WSVYSLSPHLHEDPISLDLQAGLTCCLALQUPL WSVSTEFVTHARCSLIYCWFLELAVAVQPL WSVSTEFVTHACSLIYCWFLELAVAVQPL WSVSTEFVTHACSLIYCWFLELAVAVQPL WSVSTEFVTHACSLIYCWFLELAVAVQFL WSVSTEFVTHACSLIYCWFLELAVAVQPL WSPEGGGFGAALFRENTATHHYQAMA PVYALDRINGTSRTCEFERLSTAMH PV	1			l j	ļ i		LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
KILMHETOPPSHESVSTTRIYRGTYALIFIGGE ALCLISTAFPVCIWSLGWHGGVPPLSADESR KSCTVGMATMILTLISSAWPLDLSAHODAL ILAGNILAASAPKSLESSWASEEBANPAATK QEEVWPALGDRALVPMVEQLFSHLKVNIC AHVLDDVAFOPAIRAALPSLTNPFSLSPIRKR GKEKEPGQASVPLSKGSEASARS(SDTS GPVTTSKSSLGSFYHLFSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFRSALDVLSQLEL ATIQDIGKCVEELIGYLKSCTSREPMMATVC VQULKTLFGTINLASQTDGLSNPSKSQGBA QRLGSSSVRPGLYTYCKMPSQLYTTIFTQALADA SLRNMVQAEOENDTSGWFDVLQKVSTQLKT NLTSYTKNRADKNAHHHITEPLVKALKQ YTTTTCVQLQKGVLDLLAQLVQLRVNYCLL DSDQVFIGFYLKQFEYEVGGPLSALDIVLSGGBA GREGSSVRPGLYWCMPSTERSALIPNIFF FLVLLSYERYHSKQIIGPKIQLCCOGNASGR KAVTHAPALQPIVEDLFVLAGTNKADAGKE LETQKEVVVSMLRLIQYHQVLEMFLVLQQ CHENEDKWKRLSRQIADILPMLKQGSTED VINSRQELSFSPYLISCTVINRLDGDSTSTLE BHSEGKQIKNLPEETFSFFLQLVGLILEDIVT KQLKVEMSEQUHTPVQCELGTLLMCLHIFKS GMFRRITAAATRLFSBGCGGFSTYLDSINLR ARSMITHPALAVLUWCQULLLVNHTDVRWW AEVQQTPKRHSLSSTKLLSPQMSGGEEDSILA AKLGMCNREFVRGALLILLVNHTDVRWW AEVQQTPKRHSLSSTKLLSPQMSGGEEDSILA AKLGMCNREFVRGALLILLVCONLLIKTLYQLCICL HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRAVBMLAANIQSSKAQLPBELBRIQGY LQSSGLAGRHQRUSTLOGHERLSTMQDSLSSP PVSSHPLDGDGHYSLGTLSTMCTONLOHIENS AASGLFIOAJOGRCSDLTSTTMCTTTTPFRVLARMVDIL ACRAVBMLAAANLQSSMAQLPBELBLRIQGY LQSSGLAGRHQRUSTLDIERRLSTMQDSLSSP PVSSHPLDGDGHYSLETVSPDKDWYVHLVX SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEPHLSLLAPCLSLGMSEISGGKSALFAA ARSUTLARVSGRCSINLSTFURLSTMQDSLSSP PVSSHPLDGDGHYSLETVSPDKDWYVHLVX SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEPHLSLAPCLSLGMSEISGGCGSALFAA ARSUTLARVSGRCSDLSTTMTRALAQY LVVSKLSSHLH.PPEKKEBURVAVTLEALL SWHLIHEQDFLSLDLQAGLDCCCLALQLOGL WSVVSSTEFVHACSLIYCVELAKENSOVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSRKPGGGFGTAFFEIFVEFLQKEWFKEFTYR INTLGWSRTGGGETALSHELLLUKRNSOVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSRKPGGGFGTAFFEIFVEFLQKEWFKEFTYR INTLGWSRTGGGETALSHELLLUTCHAKNSOVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSRKPGGGGGAMYSKENIATHHLYQKWD PYSLSSATTGALSHEKLLLQHAWDGHELGSWF	1		[Į į	į i	LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
VTMENNI,SRVIAAVSHELITSTIKALITGCCE ALCLISTAFPCIWSIG,GWHGPLSABBESR KSCTVGMATMILTLISSAWFFLDLSAHDOAL ILAGNILAAS,AFKSLRSSWASEEANPAATK QEEVWPALGDRALVPMYEQLFSHLLKVINIC AHVLDDVAFGPAIKAALPSLTVPPSISPIRK GKEKEPGPQASYPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFIRSALDVLSQILEL ATLQDIGKCVEEILGYLKSCTSREPMMATVC VQQLKILFGTMLASQFDGLSSNFSKQGRA QRIGSSSVRPGLYHSYCEMAPYTHFYQALADA SLERNMYQAEGEDDTISGWPVCKYSTQLKT NLTSVTKNRADRNAHHBIRLFEPLVKALKQ YTTTTCVQLQKOYDLLAGVQLVSTQLKT NLTSVTKNRADRNAHHBIRLFEPLVKALKQ YTTTTCVQLQKOYDLLAGVQLVSLTQLKT SELVELSTER STELVELY SELVELY SELVELY SELVELY ETVILSYBRYHSKQIIGFKIIQLCDGIMASGR KAVTHAPALQYPWHDLFVLKGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFLVLQQ CHKENBEKWKRISRQADIURAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLISSMF VTPNTMASVSTYQLWISGILALRVLISGSTED IVLSRIGELSFSPYLISCTVINRLRGDGSTSTLE EHSEGKQIKNLPETFSSFLIQLVGILLEDIVT KQLKVEMSEQQHTFVQQELGTLLMCLHBKS GMFRITAAATRLFRSDGCOGSFYLDSLNLR ARSMITHPALVLLWCQILLLVNHTDTRWW AEVQQTFKRHSLSSTILLSFQMSGEEGEDSLA AKLGMCNREIVBRGALLLCVVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVERNS AASGLFQAIQSGCENLSTTMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRYBMLLAANLQSSMG/SCENLSTTMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRYBMLLAANLQSSMG/PRESLYRDK AASGLFQAIQSGCENLSTTMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRYBMLAANLQSSMG/PRESLYRDKYVVHILVK SQCWTRSDSALLEGAELVNRIPAEDMAFM MNSEPNLSLAPCLSLGMSEISGGCGALIFAA AREVILARVSGTVQQLFAVHPQPELFAEP AAYWSKLNDLFGDAALYQSLFTLARALQY LVVSKLPSHLH.PPEKKEDVVAVTLEAL SWHLHEQIPLSLDLQAGLDCCCLAQLIFGL WSVVSSTEFVTHACSLIVCVHTPQFLFAEP AAYWSKLNDLFGDAALYQSLFTLARALQY LVVSKLPSHLH.PPEKEEDTLAAVQPG EQLLSPERRINTFKASEEPEEVDPHTONFKYI TAACCHAVBELAALLLPPEKEEVTVATLAALOY LVVSKLPSHLHLPPEKEVTLAAVAGVG EGSPPEEDTERTONFKAILGRKNSGVPA FLTPLLRMIISLARLPLVNSYTRVPPLVWKLG WSPKPGDDGTAFFETPVFLAGLVVAGTEALLYAMT PVAGRPAVSCLEQOPRNSHKALDTRGKK PRITCHTSMISLARLPLVNSYTRVPPLVWKLG WSPKPGDDGTAFFETPVFLAGLWCVAMD VPYSLSFATTGALSHEKLLLQHRNSGVPA FLTPLLSMISLARLPLVNSYTRVPPLVWKLG WSPKPGDGTGTAFFETWALTALANTATHRICAND PYPSLSFATTGALISHEKLLLQHAWD	i		1 1		Į i		IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
VTMENNI,SRVIAAVSHELITSTIKALITGCCE ALCLISTAFPCIWSIG,GWHGPLSABBESR KSCTVGMATMILTLISSAWFFLDLSAHDOAL ILAGNILAAS,AFKSLRSSWASEEANPAATK QEEVWPALGDRALVPMYEQLFSHLLKVINIC AHVLDDVAFGPAIKAALPSLTVPPSISPIRK GKEKEPGPQASYPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFIRSALDVLSQILEL ATLQDIGKCVEEILGYLKSCTSREPMMATVC VQQLKILFGTMLASQFDGLSSNFSKQGRA QRIGSSSVRPGLYHSYCEMAPYTHFYQALADA SLERNMYQAEGEDDTISGWPVCKYSTQLKT NLTSVTKNRADRNAHHBIRLFEPLVKALKQ YTTTTCVQLQKOYDLLAGVQLVSTQLKT NLTSVTKNRADRNAHHBIRLFEPLVKALKQ YTTTTCVQLQKOYDLLAGVQLVSLTQLKT SELVELSTER STELVELY SELVELY SELVELY SELVELY ETVILSYBRYHSKQIIGFKIIQLCDGIMASGR KAVTHAPALQYPWHDLFVLKGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFLVLQQ CHKENBEKWKRISRQADIURAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLISSMF VTPNTMASVSTYQLWISGILALRVLISGSTED IVLSRIGELSFSPYLISCTVINRLRGDGSTSTLE EHSEGKQIKNLPETFSSFLIQLVGILLEDIVT KQLKVEMSEQQHTFVQQELGTLLMCLHBKS GMFRITAAATRLFRSDGCOGSFYLDSLNLR ARSMITHPALVLLWCQILLLVNHTDTRWW AEVQQTFKRHSLSSTILLSFQMSGEEGEDSLA AKLGMCNREIVBRGALLLCVVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVERNS AASGLFQAIQSGCENLSTTMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRYBMLLAANLQSSMG/SCENLSTTMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRYBMLLAANLQSSMG/PRESLYRDK AASGLFQAIQSGCENLSTTMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRYBMLAANLQSSMG/PRESLYRDKYVVHILVK SQCWTRSDSALLEGAELVNRIPAEDMAFM MNSEPNLSLAPCLSLGMSEISGGCGALIFAA AREVILARVSGTVQQLFAVHPQPELFAEP AAYWSKLNDLFGDAALYQSLFTLARALQY LVVSKLPSHLH.PPEKKEDVVAVTLEAL SWHLHEQIPLSLDLQAGLDCCCLAQLIFGL WSVVSSTEFVTHACSLIVCVHTPQFLFAEP AAYWSKLNDLFGDAALYQSLFTLARALQY LVVSKLPSHLH.PPEKEEDTLAAVQPG EQLLSPERRINTFKASEEPEEVDPHTONFKYI TAACCHAVBELAALLLPPEKEEVTVATLAALOY LVVSKLPSHLHLPPEKEVTLAAVAGVG EGSPPEEDTERTONFKAILGRKNSGVPA FLTPLLRMIISLARLPLVNSYTRVPPLVWKLG WSPKPGDDGTAFFETPVFLAGLVVAGTEALLYAMT PVAGRPAVSCLEQOPRNSHKALDTRGKK PRITCHTSMISLARLPLVNSYTRVPPLVWKLG WSPKPGDDGTAFFETPVFLAGLWCVAMD VPYSLSFATTGALSHEKLLLQHRNSGVPA FLTPLLSMISLARLPLVNSYTRVPPLVWKLG WSPKPGDGTGTAFFETWALTALANTATHRICAND PYPSLSFATTGALISHEKLLLQHAWD					l i	l i	KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD
ALCLISTAPPUCIWSLOGWIFGOVPPLASADESR KSCTVGMATMILTLISSAWPIDLSAHODAL ILAGNILAASAPKSLRSSWASEEBANPAATK QEEVWPALGDRALVPMVEQLFSHLIKVNIC AHVLDDVAPGPAIRAALPSI.TNPPSI.SPIRKK GKEKEPGEQASVPLSPKKGSEASAASKQSDTS GPVTTSKSSI.GSFYHLPSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKUVEELIGYLKSCTEMPMATVC VQULKTLPGTNLASQPDGLSSNPSKSQGRA QRIGSSSVPBCI.YHYCPMASPMMATVC VQULKTLPGTNLASQPDGLSSNPSKSQGRA QRIGSSSVPBCI.YHYCPMASPMMATVC VQULKTLPGTNLASQPDGLSSNPSKSQGRA QRIGSSSVPBCI.YHYCPMASPMMATVC VQULKTLPGTNLASQPDGLSSNPSKSQGRA QRIGSSSVPBCI.YHYCPMASPMMATVC VQULKTLPGTNLASQPDGLSSNPSKSQGRA QRIGSSSVPBCI.YHYCPMASPMMATKLE SLENWVQAEQENDTSGWFDVLQKVSTQLKT NLTSYTKNRADKNAIHHIRLPEPLVIRALKQ YTTITTCVQLQKQVLDLLAQLVQLKVNYCLL DSSQVPIGFVLKQFEVEVGQFRESEAIPINFF FLVLLSYERYFSKQGIGFKIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLRLQIYQFULEFILLVLGHILVLQQ CHENEDBKWRRLSRQIADILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLYGLEFILVEMFLULVQQ CHENEDBKWRRLSRQIADILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLYGLUFMILVLQW VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRQIGLSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEEFISSRFLLQVGILLEDINT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRITAAATRLFSBGGCGGTTLDGVLMIPKS GMFRITAAATRLFSBGGCGTTLDSVLNIK ARSMITTIPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPGNSGGEEDSDLA AKLGMCNREIVRRGALLIPCDVYCQNLHDSE HLTWLIVNHIODLISLSHEPPVQDFISAVHRNS AASGLFQAIQSRCENLSTFTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRYEMLAANLQSSMAQPLSLNIPELNRQEY LQSGLAQRHGRLYSLLDRFRLSTMUSLSFS PPVSSHPLDGGDHVSLETVSPRDWVYVHLVK SQCWTRSDSALLEGAELVNRPAEDMAPM MNSEPLISLLARALGSSMAQDAFM MNSEPLISLLARCSLGGGRSALFEA AREVILARVSGTVQQLAVHHVPQPELPAEP AAYWSKLNDLFGDAALYGLUFALAQY LVVVSKLPSHLHIPPERKENDIVRVVATLEAL SWILLHEQIPLSLDLQAGLDCCLALQIPGL WSVVSSTEFVTHAGSLIVCVSLAAMT YVVAGNPAVSCLEQQPRNRPLKALDTRGKK LSIRGIVSGGAMYSKENLATHHLUGVAWD PVPSLSPATTGALISHEKLLLQHKNSGVPA FLTPLLKNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFEDPWETLGAMYATHLAAMT VPVAGNPAVSCLEQQPRNRPLKALDTRGKK LSIRGIVSGGAMYSKENLATHHLUGVAWD PVPSLSPATTGALISHEKLLLQHKRNSGVPA LSIRGIVSGGIAMYSKENLATHHLUGVAWD PVPSLSPATTGALISHEKLLLQHKNGVPA	1		l i				
KSCTVGMATMILTLISSAWFFLDLSAHQDAL ILAGNILAASAKSIRSISWAFSLANDATK QEEVWPALGDRALVPMVFGLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRKK GKEKEPGEQASVPLSPKKGSASAASRQSDTS GPVTTSKSSSLGSFYHLFSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFRSALDVLSQILEL ATLQDIGKCVEEILGYLKSCTSREPMAATVC VQQLKILFGRINLASQPGIGSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEGEDITSGWPVCKYSTQLKT NLTSVTKNRADRNAHHHIRLFEPLVKALKQ YTTTTCVQLQKQYDLLAGLQVLGVSTQLKT NLTSVTKNRADRNAHHHIRLFEPLVKALKQ YTTTTCVQLQKQYDLLAGLQVLGVSTQLKT NLTSVTKNRADRNAHHHIRLFEPLVKALKQ YTTTTCVQLQKQYDLLAGLQVLGVSTQLKT NLTSVTKNRADRNAHHHIRLFEPLVKALKQ YTTTTCVQLQKQYDLLAGLQVLGVSTQLKT SULVEYGFYHKQFSTEVGQFRESEAIIPNIFF FLVLLSYBRYHSKQIIGFKIQLCOGIMASGR KAVTHAPALQPVHOLFVLKGTTSKADAGKE LETQKEVVVSMLLRLQYHQVLEMFLVLQQ CHKENBEKWKRSRQADIFILAGVGMHI DSHEALGVLNTLFEILAPSSLRVDMLLRSMF VTPNTMASVSTYQLWISGIALRAVLISQSTED IVLSRIQELSFSPYLISCTVINRLRQDSTSTLE EHSEGKQIKNLPEETFSRFLQLVGILLEDIVT KQLKVEMSEQQHTFVQQELGTLLMCLHIFKS GMFRITAAATRLFRSDGGGGTYLDISLIR ARSMITTHPALVLLWCQILLLVNHTDYRW AEVQQTPKRHSLSSTILLSFQMSGEEEDISLA AKLGMCNREIVRRGALLECDVVCQNLHDSE HATWLIVNHQDLISLSHEPPVQDFISAVRINS AASGLFQAIQSGCENLSTTHKKKTLQCLEGI HSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRYBMLLAANLQSSMAQTREINKKATLQCLEGI HSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRYBMLLAANLQSSMAQTHESINNGVSPS PPVSSHPLDGGGHVSLETVSPKDWYVYHLVK SQCWTRSDSALLEGAELVNRIPAEDMAAPM MNISEPNLSLAPCLSLGMSEISGGKSALFEA ARRYTLARVSGTVQQLFAVHVPQPELPAEP AAYWSKLNDLFGGBALLYGAGLEGALVRIPSAAAMPM MNISEPNLSLAPCLSLGMSEISGGKSALFEA ARRYTLARVSGTVQQLFAVHVPQPELPAEP AAYWSKLNDLFGGDHVSLETVSPKDWYVYHLVK SQCWTRSDSALLEGAELVNRIPAEDMAAPM MNISEPNLSLAPCLSLGMSEISGGCKSALFEA ARRYTLARVSGTVQQLFAVHVPQPELPAEP AAYWSKLNDLFGGBALLYGAGLEAVAVGG EQLLSPERRITTFKASIESEEEVDPHTOMYKYI TAACCAMVAEKALGHRNSGVPA FLTPLLRMIISLARLPLVNSYTRVPPIVWLIG WSVKSTEFVTHACSLITVCVHRACALOFKNSGVPA FLTPLLRMIISLARLPLVNSYTRVPPIVWLIG WSVKSTEFVTHACSLITVCVHRACALOFKNSGVPA FLTPLLRMIISLARLPLVNSYTRVPPIVALGAMT VPVAGRPAVSCLEQQPTRINKFKALDTRGRK UNIGWTSGGGGAMYSKENIATHHLYQAWD PVPSLSARTIGALSHEKLLLORPROKADA VPVAGRPAVSCLEQQPTRINKFLALDTRGRK LSIRGIVGGGGAMYSKENIATHHLYQAWD PVPSLSARTIGAMPHORADA	1					1	
ILAGNILAASAPKSI.RSSWASEEEANPÄÄTK QEEVWPALGDRALVPWEQLFSHLIKVINIC AHVLDDVARGPAIKAALPSI.TNPPSI.SPIRKS GKEKEPGEQASYLSPKKGSEASARROSDTS GPVITSKSSSUGSFYHLPSYLKLHDVIKATHA NYKVTLDLQNSTEKFGGFI.RSALDVSQILEL ATLQDIGKCVEELGYJ.KSCFSREPMAATVC VQOLLKTLFGTNI.ASGFDGLSSNFSSOGRA QRLGSSSVRGI.HYCFMAPYTHFTQALADA SLRIMVQAEQENDTSGWEDVLQKVSTQLKT NILTSYTRANADKNAEINHLFEPLIVKALKO YTITTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVI.KQFFSYEVOQFRESEAIINHFF FILVLLSYERYHSKQIGEPKRQLCDGIMASGR KAVTHAPALQPIVHDLFVLRGTNKADAGKE LETYKEVVYSMILRILQVHQVLEMFILVLQQ CHKENEDKWKRI.SRQIADIILPMLAKQOMHI DSHEALGVINTLFELLASRVYDMLRSMF VTPNTMASVSTVQLWISGILAILRVLISGSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTILE EHSEGKQIKNLPEETFSRFLLQLVGILLEDINT KQLKVEMSEQQHTTYCQELGTLIMCHIEKS GMFRRITAAATRLFRSDGCGGSFYTLDSINLR ARSMITTHPALVLLWCQILLIVINTDYRWW AEVQQTFKRISLSTKILSFQMSGGEEDSDLA AKLGMCNREIVRRGALLIFCDYVCONLHDSE HLTWLIVNHIQDLIS.SHEPPYODFISAVIRNS AASGLFQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACKRVEMILAANLQSSMAQLPMEEINRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPYSSHPLJGDGOHYSLETYSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEFNLSLLAPCLSLGRANQLPMEEINRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMODSLSPS PPYSSHPLJGDGOHYSLETYSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEFNLSLAPCLSLGRASGGGGSGGKALFEA AREVILARNVSGTVQQLPAVHHVPGPELPAEP AAYWSKLDDGOHYSLETYSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEFNLSLAPCLSLGRASGGGKSALFEA AREVILARNVSGTVQQLPAVHHVPGPELPAEP AAYWSKLDDGOHYSLETYSPDKDWYVHLVX SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEFNSLSLAPCLSLGRASGGGKSALFEA AREVILARNSGTVQQLPAVHHVPGPLPAEP AAYWSKLDDGOHYSLETYSPDKDWYVHLVX SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEFNSLSLAPCLSLGRAGLIVGRPAGERYPY TAACEMWAEMSLGSVLLLAGHRINSGVYI TAACEMWAEMSLGSVLLLAGHRINSGVYI TAACEMWAEMSLGSVLLLAGHRINSGVYI TAACEMVAEMVESLGSVRALLAGHRINSGVYI TAACEMVAEMVESLGSVRAPLALAGHTUNGAVING SWELPGGGTAPPEDEFGLEEVEVFREFIY NTLGWTSRTQFEETWATLLQVLVTQPLVMLE QEESPPEEDTERTQNVLAVQAGTGSLKSMI	1	Į į	1 I	ļ :		ļ	1
QEEWPALGDRALLYMVREQLESHLIKEVINIC AHVLDDVAGPAIKAALPSILTMPSILSPIRK GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSSYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKTGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSGFSEFMATTVC VQOLLKTLGTNLASGPTOLSSNPSKSOGRA QRLGSSSVRPGLYHYGFMAPTYTHFTQALADA SLRMWQAGENDTSGWDVLGKYSTQLKT NLTSVTKNRADKNAHNHRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRYNYCLL DSDQVTIGFVLKQFFSYEVGOFRESEAHIPNFF FLVLLSYERYHSKQIGIPKHQLCDGIMASGR KAVTHAPPALQPVHDLFVGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDK WKRLSRQIGIPKHQLCDGIMASGR KAVTHAPPALQPVHDLFWATNAGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDK WKRLSRQIGHEMILVALQSTED IVLSRIGELSFSYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILEDIVT KQLXVEMSEQQHTFYCGLGTLLMCLHIFKS GMFRRITAAATRLFRSDGGGGFYTLDSLNLR ARSMITTHALVLLWCQILLDIVT KQLXVEMSEQQHTFYCGLGTLLMCLHIFKS GMFRRITAAATRLFRSDGGGGFYTLDSLNLR ARSMITTHALVLLWCQILLDHIVT KQLXVEMSEQQHTFYCGLGTLLMCLHIFKS GMFRRITAAATRLFRSDGGGGFYTLDSLNLR ARSMITHALVLLWCQILLDHIVT KQLXVEMICAATRLIJCHTPRVLARMVGIL AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISSIBEPPVQDFISAVHRNS AASGLFQAQQSRCENLSTPTMLKTLQCLEGI HLSOSGAVILTYVDRLLCTFRVLARMVGIL ACREVEMILAANLQSSMAQLPMEEINRIQEY LQSSGLAQRRQLYSLLDGTSLTSMQDSLYFNS AASGLFQAQQSRCENLSTPTMLKTLQCLEGI HLSOSGAVILTYVDRLLCTFRVLARMVGIL ACREVEMILAANLQSSMAQLPMEEINRIQEY LQSSGLAQRRQLYSLLDGTLSTMQDSLAFA AREVILARVSGTVQQLPAVHHYPQELPAEP AAYWSKLDJLFAGALQSGGGKSALFFA AREVILARVSGTVQQLPAVHHYPQELPAEP AAYWSKLDJLFDAALVGFILARALQQY LVVYSKLPSHIHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLCCCLALQLPGL WSVYSSTEVTHACSLTYCVHFILEAVAVQPG EQLLSFERTTNITKAISEEEEEVDPINQNRKYI TAACEMVAEMVESLGGVALLGIKKNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSVYSSTEVTHACSLTYCVHFILEAVAVQPG EQLLSFERTTNITKAISEEEEEEVDPINQNRKYI TAACEMVAEMVESLGGAVKPLAALDTRGRK NTLGWTSRTOPEETWATLLQULVTOPLVME QEESPPEEDTERTQINVLAVQALAGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSVSSTEVTHACSLTYCVHFILEAVALVPGL	1	Į į	1 I	ļ ¹	l i	Į l	
AHVLDDVAPGBAIKAALPSLTNIPPSLSPIRIK GKEKEPGGQASVPLESKUSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVSQILEL ATLQDIGKCVEELGYLKSGFSREPMMATVC VQQLLKTLFGTNLASGFDGLSSNFSSGGRA QRLGSSSVRGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKAIHNHIRLFPLVKTALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFFYEVGGFRESEAIPNIFF FLVLLSYSERYHSKQIGIFPKIQCLCDGIMASGR KAVTHAPALQPVYDLFVLRGTNKADAGKE LETOKEVVVSMLLRLIDYOLEMFILVLQQ CHKENEDK WKRLSRQIADILIPMLAKQQMH DSHEALGVLNTLFEILAPSSLRPVDMLRSMF VTPNTMASSVTVQLWISGLAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQKINN PEETFSRLQVGGLILMCLHIFKS GMFRITAAATRLFRSDCCGGSFYTLDSINLR ARSMITTHAVLJLWCQLLLVHITDYRWW AEVQQTFKRISLSSTKLLSPQMSGEEDSDLA AKLGMCNERVFRGALLISPQMSGEEDSDLA AKLGMCNERVFRGALLISPQTWGSEEDSDLA AKLGMCNERVFRGALLISPQTWGSEEDSDLA AKLGMCNERVFRGALLISPQTWGSLEEDSDLA AKLGMCNERVFRGALLISPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARNYDIL ACREVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAGRIQRLYSLLDFRLSTMODSLSPS PPVSSIPLDGDGHVSLETVSFDKDWYHLVK SQCWTRSDSALLEGABLVRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTYQQLFAVHHYPGELPAEP AAYWSKLNDLFGDAALVGSLPTLARALAQY LVVVSKLPSHLHLPPEKKENDVKFVATLEAL SWHLIFEGPISLDLOAGLOCCLALQPGL WSVVSSTEFVTHACSLIYCVHFILBAVAVQPG EQLLSPERTNTIPKAISEEEEVPIPNTONPKYI TAACEMVAEMWESLQSVLALGHERNSGVPA FLTPLLRNIISLARLPLVNSYTRVPFLVWALG WSVPSGGFGTAFFEEFFULPKEVFIYW NTLGWTSKTOPEEF WATLLGVLNTPLWALG WSVPSGGFGTAFFEEFFULPKEVFIYW NTLGWTSKTOPEEF WATLLGVLNTPLWALG WSVPSGGFGTAFFEEFFULPKEVFIYW NTLGWTSKTOPEEFWATLLGVLNTPLWALG WSVPSGGFGTAFFEEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOPEFFULPKEVFIYW NTLGWTSKTOTHETHALY NTLGWTSKTOPEFFULPKEVFIYM NTLGWTSKTOTHETHALY NTLGWTSKTOTHETHALY NCHANCOTHALY NTLGWTSKTOTHETHALY NTLGWTSKTOTHALY NTLGWTSKTOTHALY NTLGWTSKTOTHALY NTLGWTSKTOTHALY NTLGWTSTOTHALY NTLGWTSTOTHALY NTLGWTSTOTHALY NTLGWTSTOTHALY NTLGWTSTOTHALY NTLGWTSTOTHALY NTLGWTSTOT	1		1)	, ,	Į i	1	1
GKEKEPGGASVELSENKIGSEASAASRQSDTS GPVTTSKSSLGSFYHLSVLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCFSREPMATVC VQQLLKTLFATNLASQFDGLSSNPSKSQGRA QRLGSSSVRFGLYHYCFMAPYTHFTQALADA SLRMVQAGENDTISGWDLGKVSTQLKT NLTSVTKNRADKNAHNHRLFEPLVKALKQ YTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFYEVEVGPFRESAHPNIFF FLVLLSYERYHSKQIGIFKIQLCDGMASGR KAVTHAPPALQPVHDLPVLRGTNKADAGKE LETQKEVVVSMLIRLIQVYHQVLEMFLVLQQ CHKENEDKWKRLSRQIADILPMLAKQQMHI DSHEALGVINTLFEILAPSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED IVLSRQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKM.PEETFSRFLLQLVGILEDIVT KQLKVEMSEQQHTFYCQLGTLLMCLHIFKS GMFRRITAAATRLFRSDGGGGSFYTLDSLNLR ARSMITTHALVLLWCQILLLVHITDYRWW AEVQQTFKRISLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCORLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVYRNS AASGLFQAQSRCENLSTPIMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACREVEMLAANLQSSNAQLPMEELNRIQEY LQSSGLAQRIQRLYSLLCTFFRVLARMYDIL ACREVEMLLAANLQSSNAQLPMEELNRIQEY LQSSGLAQRIQRLSTLTTYDRLKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACREVEMLLAANLQSSNAQLPMEELNRIQEY LQSSGLAQRIQRLSTLDGTSTLSTMQDSLTARALAQCY LVVSKLPSHLLBAPCLSTLARDGASAGLFA AREVTLARVSGTVQQLPAVHHYPQFELPAEP AAYWSKLDDFDAALDSTLARDMAPM MNSEFNLSLAPCLSLAFUSPLAEDMAPM MNSEFNLSLLAPCLSLOFTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLDLQAGLUSTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLDLQAGLUSTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLSLDLQAGLUSTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLDLQAGLUSTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLDLQAGLUSTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLDLQAGLUSTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLDLQAGLUSTLARRLAQQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIBEQPISLDLQAGLUSTLARRLAQGYPA PLTPLLRNIBLSLARPLVNSYTRVPPLVWKLG GESPPEEDTERTQINVLAVQATISLVLSAMT VPVAGNPAVSCLEQQFRNKFLKALDTRFGRK LSIRGIVSQEIQAMVSKRENLATHHLYQAWD PVPSLSPATTOALSISLKLLQNPTERELGSMS	1		1	ı l	l i	ļ	l
GPVTTSKSSLGSFYHLPSYLKLHDVLKATHAA NYKVTLDLQNSTEKFGERSALDVLSQLEL ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTTNLASQFPGLSSNFSKSQGRA QRLGSSSYRGQLYHYCMAPYTHFTQALADA SLRINMYQAEQENDTSGWFDVLOKYSTQLKT NLTSYTKINRADKNAHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYTEVGQFRESEAILPNIFF FLVLLSVERYHSKQIGDFKIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFIL VLQQ CHKENEDK WKRLSRQIADDILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAURVLISQSTED IVLSRIQELSFSPVLISCTVINRLINGDENSTSILE EHSEGKQKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRRTAAATRLFSSTRLLLSQMSGEEDSDLA AKLGMCNEIVRRGALIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAJQSRCENJFTMLKKTLQCEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLAANLQSSMQLPFMELNRIQEY LQSSGLAGRIGRLYSTPMLKKTLQCEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLAANLQSSMQLPMELNRIQEY LQSSGLAGRIGRLYSLTMRPAEDMAAFM MNSEPILSLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALQSLPTLARALAQY LVVVSKLPSH-HLPPEKEEDIVKFVVATLEAL SWHLHERDPISLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLLYCVHTILEAVAQPQG EQLLSPERRTNTPKAISEEEEVDPNTQNYKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLIFLLRNIIISLARLP.VNSYTRYPLVWKLG GESPPEEDTERTONVLAVQATISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGKK LSIRGIVEQEIQAMYSKRBNATHHLYQAWD PVYSLSPATTGALISHKLLQNFERELGSMS		ļ Ì	! I	ļ l	ļ l		l
NYXVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRFGLYHYCFMAPYTHFTQALADA SLRINMYQAEQENDTSGMPVDLQKVTQLXTDAA SLRINMYQAEQENDTSGMPVDLQKVTQLXTDAA SLRINMYQAEQENDTSGMPVDLQKVTQLXTDAA SLRINMYQAEQENDTSGMPVDLQKVTQLXALAQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVGQFRESEAUPNIFF FLVLLSYERYHSKQIIGFHKIQLCDGIMASGR KAVTHAPALQPVHDLFVLRGTNKADAGKE LETQKEVVVSMLLR.I.QYHQVLEMFIL.VLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHBALGVLNTLFFILAFSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILALIRVLISQSTED IVLSRIQGLSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEFTSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLUWCQULLLVHTIDYRWW AEVQQTPKRHSLSSTKLJSPQMSGEEDSDLA AKLGMCNREVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRVEMLAANLQSSMAQLPMEELINIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRTAEDMAFM MYSEPNLSLLAPCLS.IGMSEISGQKSALFEA AREVTLARVSGTVQQLPAVHHYFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHHLPFEKERDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPSRTNTPKALSEEPEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITLILRNIIISLARLPLVNSYTRYPLVWKLG SWSPKRGGDFGTAFFEPVEFTQEKEVKFEFTYR INTLGWTSRTQFEETWATLLGVLYTOPLVME QEESPPEEDTERTQNVLAVAQATISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIYEQEIQAMYSKRENNATTHLYQAWD PVYSLSPATTGALISHEKLLLQNNFERLEGMS	1.		۱ I	į l	ı l		1
ATLQDIGKCVEELIGYLKSCFSREPMATVC VQQLLKTLFGTNLASQFGLSSPPSLSQGRA QRLGSSSVRPGLYHYCPMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSQVVFIGFYLKQEPVGQFRESEALPINFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFIL VLQQ CHKENEDKWKLSRQIADILHASQLRHHI DSHEALGVNTLFEILAPSSLRPVDMLLSSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSPSYLISCTVINRLINGDSTSTILE EHSSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTILMCLHIFKS GMFRRITAAATLFRSDGCGGSFYTLDSILNC ARSMITHPALVLLWCQULLLVHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILPCDYVCQNLHDSE HLTWLIVNHIQDLISLSSEPPVODFISAVHRNS AASGLFJQAQSRCENTJEMLKKTLQLEGI LSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHGRLVSSLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGELVNRTPAEDMAFM MISSENISLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALQY LVVVSKLPSH-HLPPEKEEDIVKFVATLAALAQ VLVVSKLPSH-HLPPEKEEDIVKFVATLAALQY LVVVSKLPSH-HLPPEKEEDIVKFVATLAALQ SWHLHERPISLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNITKALSEEEVDDPNTONPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITLLRNIIISLARLPLVNSYTRYPLVWKLG WSPKFRGGOFGTAFFEPVEFFLQEKEVFKEFIYR NTLGWTSRTOFFETWATLLGVLVTOPLVME QEESPFEEDTERTQINVLAVQATISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEIQAMVSKRBNATHHLYQAWD PVYSLSPATTGALISLEGMS	1		ļ Ì			ļ	[
VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAIENHERLFEPLVIKALKQ YTTTTCVQLQKQVLLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVGGGRESEAHPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIYHDLFVLRGTNRADAGKE LETQKEVVVSMLLRLQYYHQLEMFILVLQQ CHKENEDKWRLISQIADIILPALAKQQMH DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED IVLSRIQELSFSPYLJSCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AFVQQTPKRHSLSSTLSTVMSLRDGMSGEEEDSDLA AKLGMCNREIVRGALLFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVFRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYYDRLLCTPFRVLARMVDIL ACRRVEMLLAANLSSMAQLFMEELINGQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVX SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLJAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQLFAVHHVPOFELPAEP AAYWSKLNDLFGDAALYQSLFTLARALQY LVVVSKLPSHLHPPEKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVFPILLARALQY LVVVSKLPSHLHPPEKEKDIVKFVVATLEAL SWHLHEGPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVFPILLARALQY LVVVSKLPSHLHPPEKEKDIVKFVVATLEAL SWHLHEGPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVFPILLARALQY LVVVSKLPSHLRIPRASEEEEEVDPINQPKYI TAACEMVAEMYESLQSVLALGHKRNSGVPA FLITPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEFVEFLQEKEVKEFTYN INTLGWTSRTQFEETWATLLGVLVTQPLVME GESPPEEDTERTQIVALQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVFPILLEAVAVQPG EQLLSPERTTITFRASEEEEEEVDPINQPKYI TAACEMVAEMYESLQSVLALGHKRNSGVPA FLITPLLRNIIISLARLPLVNSYTRVPPLVWYLG WSPKPGGDFGTAFPEFVEFLQEKEVKEFTYN INTLGWTSRTQFEETWATLLGVLVTQPLVME GESPPEEDTERTQIVALAQATISLJLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEIQAMVSKERNIATHHLYQAWD PVPSLSPATTOTALSHEKLLLQNPTERLGSMS			ļ l		ļ l	ļ	
QR.GSSSVPGLYHYCFMAPYTHETQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAEHNHRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFSYFEVGGPRESEAHPNIFF FLVILSYERYHSKQIGGFKIIQLCDGMASGR KAVTHABALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFALFSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRQLEISFSPYLISCTVINBLRDGDSTSTLE BHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHHEKS GMFRRTAAATRLFRSDGCGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVMHTDYRWW ABVQQTPKRHSLSSTKLLSPOMSGEEEDSDLA AALLGMCREIVRRGALIFCDYVQDNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDLLCTPFFVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLLOGDHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRIPAEDMNAFM MNSETPLSLLAPGLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALVGSIFTLARALAQY LVVVSKLPSHLHILPPEKKEDIVKFVVATLEAL SWILHEQPLSLDLQAGLDCCLALQLPGI WSVVSSTEPTHACSLIYCVHFILEAVAVOPG EQLLSPERRTNITFKAISEEEEEEVDPNTQNPKYI TAACENVAEMVESLQSVLAGHKRNSGVPA FLTPLLRMIISLARPLVNSTTRVPFILAWAVOPG GULSPERRTNITFKAISEEEEEEVDPNTQNPKYI TAACENVAEMVESLQSVLAGHKRNSGVPA FLTPLLRMIISLARPLVNSTTRVPFILAWAVOPG WSPKPGGDFGTAFPEIPVEFLQEKEVFKETYR INTLGWTSRTOPEETWAILLGVLVTQPILVME QUESPPEEDTERTQNVLAQAMTSLVLSAMT VPVAGRPAVSCLEQQPRNEPLKALDTRFGRK LSIRGIVEQEIQAMVYSKRENIATHHLYQAWD PVPSLSPATTOALSHEKLLLQNPPERELGSMS		(l	1 1	r i	1 1	ļ İ	
SIRNMYQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHNHRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFFYIEVGGFRESEAIIPNIFF FLVLLSYERYHSKQIGFKIQLCOGIMASGR KAVTHAPALQPWIDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKLLSQAIDILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VYPNTMASVSTVQLWISGILAILEVLIQSTEED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSPFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLUCOLLLVHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYYHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEPNLSLLAFCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLFTLARALAQY LVVVSKLPSHLHIGLEVEKEDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLLTVCHFILEAVAVQPG EQLLSPERRTNTFRAISEEEEEVDPNTQNFKYI TAACEMVAEMYESLQSVLALGHKRNSGYPA FILTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSVRSTEFVTHACSLLTVSTYPPLVWKLG WSVRDGDFGTAFPEFVEFLQEKEVKFFTYY INTLGWTSRTQFEETWATLLGVLVTQPLVME QESSPPEEDTERTQNVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVYSKERALTHLTVQAWD PVPSLSPATTGALSHEKLLLQNIPERLGSMS	1		1 1		ļ İ		
NLTSVTKNRADKNAHNHIRLFEPLVIKÁLKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVGQFRESEAIIPNIFF FF.VLLSYERYHSKQIIGIPKIQLCOGIMASGR KAVTHAPALQPVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQLADILLPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRFVDMLLRSMF VTPNTMASVSTVQLWISGI AILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATILFRSDGCGGSFYTLDSLIKLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYCONLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDFRRLSTMQDSLSPS PPVSSHPILDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAFSLSGGWSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALVQSLPTLARALAQY LVVVSKLPSHLHPPEKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCLALQLPGL WSVVSSTEFVTHACSLLYCVHFILEAVAVQPG EQLLSPERTNITYKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLALPLVNSYTRYPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRIQFEETWAHLGVLVTQPLVME QEESSPEEDTERTQINVLAVQAJTSLVLSAMT VPVAGNPAVSCLEQQPRKKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	1		1	·		
YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGPKIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVINTLFEILAPSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLIQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQUKNLPEETFSRFLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLIMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQULLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFQAQISGREENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEEINRIQEY LQSSGLAQRRIQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGKSALFEA AREVTLARVSGTVQCJPAVHHVPQPELPAEP AAYWSKLNDLFGDAALVOSLPTLARALAQV LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCLALQLPGL WSVVSSTEFVTHACSLYCVHIPILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLGSVLALGHKRNSGVPA FLTPLLRNIEISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFFEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVMEG QEESSPEEDTERTQINVLAVQANTSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1		1 1	ļ l	1	ļ	
DSDQVFIGFVLKQFEYEVGGFRESEAIIPNIFF FLVLLSYERYHSKQIGIPKIIQLCDGIMASGR KAVTHAPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQVHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLWCQILLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYYDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELNNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVPOPELPAEP AAYWSKLNDLPGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGIL WSVVSSTEFVTHACSLTYCHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLLGVLHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLLGVLHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLLGVLAGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKEFTYR INTILGWTSRTQFFETTWATILLGVLVTQPLVME QEESPPEEDTERTWATILLGVLVTQPLVME QEESPPEEDTERTQNIVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGAALISHEKLLLQINPERELGSMS	1	1 1	1 1		! 		
FLVLLSYERYHSKQIIGPKIIQLCDGIMASGR KAVTHAPALQPIYHDLFVLRGTNKADAGKE LETQKEVVYSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASYSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQHTFYCQELGTILMCLHIFKS GMFRITAAATRLFSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTVLINVHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVITLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEEINRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSIPLDGDGHVSLTVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVPQFELPAEP AAYWSKLNDLFGDAALYQSLTTLARALAQY LVVVSKLPSHLHIPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFTVHACSLIYCVHTILEAVAVQFG EQLLSPRRTINTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLLQSVLALGHKRNSGVPA FILTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEPVEFLQEKEVFKEFTVR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PYPSLSPATTGALSHEKLLLQINPERELGSMS			1	,	: I		
KAVTHAPALQPIVHOLFVLRGTINKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQL WISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTILE EHSEGKQIKNLPETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYQELGTILMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNRETVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIQSRCENLSTITMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSIPLDGGGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSIGMSEISGGQKSALFEA AREVTLARVGTVQQLFAVHHVPQPELPAP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEDIVKFVVATLEAL SWHLHEQPISLDLQAGLDCCCLALQLPGI WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTINTFKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FILTPLLRNIISLARPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEPVEFLQEKEVFKEFTYR INTILGWTSRTOFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALSHEKLLLQNPERELGSMS			1	1		1	
KAVTHAPALQPIVHOLFVLRGTINKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQL WISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTILE EHSEGKQIKNLPETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYQELGTILMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNRETVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIQSRCENLSTITMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSIPLDGGGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSIGMSEISGGQKSALFEA AREVTLARVGTVQQLFAVHHVPQPELPAP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEDIVKFVVATLEAL SWHLHEQPISLDLQAGLDCCCLALQLPGI WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTINTFKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FILTPLLRNIISLARPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEPVEFLQEKEVFKEFTYR INTILGWTSRTOFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALSHEKLLLQNPERELGSMS		<u> </u>					
LETQKEVVVSMLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVINTLFEILAPSLRPVDMLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQKINLPEETFSRFILQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANIQSSMAQLPMEELNRIQEY LQSSGLAQFRQLYSLLDFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIFAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQRG EQLLSPERRTNTPKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFFETWATLLGVLYTQLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSFATTGALISSHEKLLLQNPERRELGSMS			1 I	!	! 		KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIGELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRAVEMILLANNLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPFGKKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENLATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPFERLGSMS]						LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ
DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIGELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRAVEMILLANNLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPFGKKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENLATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPFERLGSMS	1		'				CHKENEDKWKRLSRQIADIILPMLAKQQMHI
VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQKKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLWCQILLLVNHITDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEGDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFQAVITLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLFAVHHVFQPELFAEP AAYWSKLNDLFGDAALYGSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFYTHACSLIYCVHFILEAVAVQPG EQULSPSFRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITLLRNIIISLARLFLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR NTILGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRIKRLKALDTRFGRK LSIRGIVEQEIQAMVSKRENLATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1		'			' <u> </u>	l
IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVPQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEVDPNTOMPK YI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR NTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			' 			'	
EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTILMCLIHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWLNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSIPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIGEY LQSSGLAQRHORLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRIPAEDMNAFM MNSEFNLSLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLAIQLPGL WSVVSSTEFVTHACSLIVCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFFEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			' I		l	' 	
KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPILDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEPNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR NTILGWTSRTQFETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1		' I	1 1	ı	' 	
GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITIHPALVILWCQILLLVNHITDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	}	ļ		1	·	' j	
ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGALILFDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			·	1	'		
AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANIQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHI.PPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHI.YQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			l		ı İ	ļ	
AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			Í				
HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			l			' !	
AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTILGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1				' l	' I	
HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		' [- 1	İ	1		
ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		· [į	· [·		HI GOGGAVI TI VVMDI I CTDEDIR ADAGOM
LQSSGLAQRHQRLYŠLLDRFRLSTMQDSLŠPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			j	· [
PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTILGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	· [1	. 1	1		
SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLPGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTILGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		· [İ	'			
MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		· J	J	1	}	ļ	
AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1 1	' I		1	ļ		
AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			l	1	1		
LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			1	İ	ļ		
SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	į l	J	J	·	I	j	
SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		l	l	-	l	İ	
WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	l			į	l	
EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1		ŀ	i	į	1	
TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	ı	J	ŀ	I		
FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHILYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	;	J	ļ	1		
WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1 1	ì	ł	ł	į.	ļ	• • • • • • • • • • • • • • • • • • • •
INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1		- 1	ł	İ	I	. — —
QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		٠ ا	- 1	I	ĺ	<u>.</u> I	
VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		·	1	ļ	İ	I	
LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1		- 1		İ	1	
PVPSLSPATTGALISHEKLLLQINPERELGSMS	1 1		- 1			.	I SID CIVE OF CALACTER TO THE COLOR OF CALACTER TO THE COLOR OF CALACTER TO THE COLOR OF CALACTER TO THE COLOR OF CALACTER TO THE COLOR OF CALACTER TO THE COLOR OF CALACTER TO THE CALACTER T
	1 1	ľ	1	4	ł	I	
YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE						ţ	
	L	l					YKLGQVSIHSVWLGNSITPLREEEWDEEEEE

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of, peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV KLSVDRVNVHSPHRAMAALGLMLTCMYTG KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM NKVIGEFLSNQQPYPQFMATVVYKVPQTLHS TGQSSMVRDWVMLSLSNFTQRAPVAMATWS LSCFFVSASTSPWVAAILPHVISRMGKLEQVD VNLFCLVATDFYRHQIEEELDRRAFQSVLEV VAAPGSPYHRLLTCLRNVHKVTTC
428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG LPCVGDAAEYQDCNPQACPVRGAWSCWTS WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL GLHTEEALCATQACPEGWS
429	1779	Ā	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/ RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI PPES/RS*QGGTVQTGQHSSGREAGSWRARGR NAGRR*KGGGKIGTKQGAVRARKECRGEMA SGETDSE
430		A	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE ASKCNIVCTQPRRISAVSLANRVCDELGCENG PGGRNSLCGYQIRMESRACESTRLLYCTTGV LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS DFLLIILKEILQKRSDLHILIMSATVDSEKFST YFTHCPILRISGRSYPVEVFHLEDIIEETGFVLE KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI PDVVFVIDTGRTKENKYHESSQMSSLVETFVS KASALQRQGRAGRVRDGFCFRMYTRERFEG FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF LSKALDPPOLQVISNAMNILIRKIGACELNEPK LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP VATLAAVMTEKSPFTTPIGRKDEADLAKSAL AMADSDHLTIYNAYLGWKKARQEGGYRSEI TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA GLYDNVGKIYTKSVDVTEKLACIVETAQGK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK MSLENDKILQIITELIKTENN
431	1781	A	3474	1	441	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ QRGEASTGGASGRRCGSCFQV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first arnino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
						QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE CSLFNKWC/YKNWM/LHVQKKRI*VQTLHPS QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY SRALNSGSR
433	1783	A	3504	1876	552	CLAPCSPQPEKNGMQPLLLLPPLLYQQLLHS SLGAPGESTLLVRTSKLLVGLGLQLLVWLLL QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS APRSRCVARPAARTGLPTPAPASSPAPAASPA PAASPAPAESTAVPQPLILLPKP/PPAPGAPPPRP GAPPPRAASPSPAASPAPPAASPVLTASPPLP AASPSPAASPAPPAASPVLTASPPLPAASPSPA ASPAPPAASPVLTASPPLPAASPALAASPVHT ASPPVHVASPPVHTASPPVHVAS
434	1784	A	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA SFFVFLV*TGF\TALARMVLISWPCDLPTSASQ SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV RTKFGINMVTSRERGTTRLPKEG
435	1785	A	3529		3161	MSLVRAALEALDELDLFGVKGGPQSVIHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC EWPLFWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTFTPOFIDQQOFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SSEFLTISCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIEIHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGIMTYLSKAAATY VQEFLPHSICAMQ
436	1786	A	3546	73	393	CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL EQPDSCRPYGRSFYALEEKHVIFSLDVGETDN

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence]	09/496	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
nence	ļ	l	914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
	\			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	'		}	peptide	Sequence	/=possible nucleotide deletion, \=possible
1	ľ	i	1	sequence		nucleotide insertion
<u> </u>			 	sequence		KGKGKTIRGI*TFKGRKGGTYQREHDANPLA
]]			PXSARSCWMRKG
437	1787	A	3554	5157	2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAOG
""	1,0.	••	300.		2,33	SASGCRYAAHPHWGLGGAAAAGGSWEPQPP
				·	i	RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA
ļ						AQPGARARTSPPPASARNMAARPAATLAWSL
			i			LLLSSALLREGCRARFVAERDSEDDGEEPVVF
					ļ	PESPLQSPTVLVAVLARNAAHTLPHFLGCLER
1					·	LDYPKSRMAIWAATDHNVDNTTEIFREWLK
				•		NVQRLYHYVEWRPMDEPESYPDEIGPKHWP
1						TSRFAHVMKLRQAALRTAREKWSDYILFIDV
					<u> </u>	DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS
1						NFWCGITPKGFYKRTPDY\VQIREWKRTGCFP
1					·	VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW
						TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP
[LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ
						YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD
1						RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA
1				i i	·	LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV
						WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK LMDNIDOAOLDWELIYIGRKRMOVKEPEKA
1						VPNVANLVEADYSYWTLGYVISLEGAQKLV
						GANPFGKMLPVDEFLPVMYNKHPVAEYKEY
						YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE
				*	·	TSTIWDNETVATDWDRTHAWKSRKQSRIYSN
						AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C
						NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA
1						NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA
						SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT
						TLG
439	1789	Α	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLLSTLTGRSY
1						GQPSLQDELKDNTTVFTRILDRLLDGYDNRL
						RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI
						DVFFRQSWKDERLKFKGPMTVLRLNNLMAS
						KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE
		· .]	DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH
						ACPLKFGSYAYTRAEVVYEWTREPARSVVV
						AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV
						MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF
						WLNRESVPARTVFGVTTVLTMTTLSISARNSL
			,			PKVAYATAMDWFIAVCYAFVFSALIEFATVN YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN
						NTYAPTATSYTPNLARGDPGLATIAKSATIEP KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL
]	.				l	FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	A	3568	1	350	STSSCFPAAAAAIMREIVHLOAGOCGNOIGAK
	1170	^	3300	•	220	FWEVISDEHGIDPTGTYHGDSDLQLERINVYY
						NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG
						GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569	2	1751	FVAVAGAVSGEPLVHWCTOOLRKTFGLDVS
***	1,71	^	3303	-	1/31	EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGO
						FIEELITKWQKNDQELISDPLQQCFKKDEILDG
				ļ		QKSGDHLKRGRKKGRNRQEVPAFTEPDITAE
				į		VKTPFDLAKAQENSNSVKKKTKFVNLYTREG
						QDRLAVLLPGRHPCDCLGQKHKLINNCLICG
			Ì	l	ŀ	RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS
				ļ		N/KSQKLLKKLMSGVENSGKVDISTKDLLPH
						QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI
						

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DDESDYFASDSNQWLSKLERETLQKREEELR ELRHASRLSKKVTIDFAGRKILEEENSLAEYH
			·			SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL VNPNMYQSPPQWVDHTGAASQKKAFRSSGF GLEFNSFQHQLRIQDQEFQEGFDGGWCLSVH QPWASLLVRGIKRVEGRSWYTPHRGRLWIAA TAKKPSPQEVSELQATYRLLRGKDVEFPNDY PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH QGAKKGLMKQNKAV
442	1792	A	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK KTAGVKPYECTICGKAFMRLSSLTRHMRSHT AIRAINEKPYKCKEC\GRAFFSLSQILSK\HERSH TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK PYECKQCGKALSCSSSLRVHERIHTGEKPYEC KQCGKAFSCSSSIRVHERIHTGEKPYACK\EC GKAFIS\TTSVLTHMITHNGDRPYKCKECGKA FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS TSIQIHERIHTGEKPYKCKECGKSFSARPAFRV HVRVHTGEKPYKCKECGKAFSRISYFRIHERT HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG EKPYECKECAKTFISLENFRRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ CGKAFSCSSYIRIHKRTHTGEKPYECKECGK AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS LHSSFR\HTRIHNYEKPLEC*Q\CGKAFSVSTS LKKPMRNAQSDRKLY\KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHVSVSNGPYR\CGSPIRLYNT*NISINRNL VAVVTP*CSTLFKCLWCWCKRAALSVV*\IVQ DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY
443	1793	A	3578	287	114	DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335		HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETQFYTAETVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK
445	1795	A	3584	1	6169	RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA FPVAFLETHLDKHNIYSIYNTKSSRERAALSLP TNVEDVCPNIPSLEKLMEEIVELAESGIRYTQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	1=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence] .	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	·	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
						MPHVMEVILPMLCSYMSRWWEHGPENNPER
1						AEMCCTALNSEHMNTLLGNILKIIYNNLGIDE
1						GAWMKRLAVFSQPIINKVKPQLLKTHFLPLM
1						EKLKKKAATVVSEEDHLKAEARGDMSEAEL
						LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL
						KEPNPEAEELFRMVAEVFIYWSKSHNFKREE
						QNFVVQNEINNMSFLITDTKSKMSKAAVSDO
1 1		1		:		ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL
1						NICAPGDQELIALAKNRFSLKDTEDEVRDIIRS
1						NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS
1						DPEKTVERVLDIANVLFHLEQKSKRVGRRHY
						CLVEHPORSKKAVWHKLLSKORKRAVVACF
						RMAPLYNLPRHRAVNLFLQGYEKSWIETEEH
						YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHO
1						LILLFSRTALTEKCKLEEDFLYMAYADIMAKS
[CHDEEDDDGEEEVKSFEEKEMEKOKLLYOO
1			ł			ARLHDRGAAEMVLQTISASKGETGPMVAAT
1						LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF
i i						FQSLAGLMQSCSVLDLNAFERQNKAEGLGM
1						VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH
						NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVO
1 1						ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA
1 1						KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV
1						VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ
						KDMVVMLLSMLEGNVVNGTIGKOMVDMLV
1 [ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD
						PDGKGVIFKRDFHKAMESHKHYTQSETEFLL
1 1						SCAETDENETLDYEEFVKRFHEPAKDIGFNVA
i l						VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP
1	ĺ		[ĺ	FLGRIEIMGSAKRIERVYFEISESSRTQWEKPO
		ĺ	j			VKESKRQFIFDVVNEGGEKEKMELFVNFCED
						TIFEMQLAAQISESDLNERSANKEESEKERPEE
1 1			j			QGPRMAFFSILTVRSALFALRYNILTLMRMLS
			•			LKSLKKOMKKVKKMTVKDMVTAFFSSYWSI
1	(1		l	FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA
		ı. İ				KKIKVAELLANMPDPTQDEVRGDGEEGERKP
]		1	1			LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR
))		j			,	EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE
1 1						KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE
	1	1	ĺ			KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA
		l				FWKKIIAYQQKLLNYFARNFYNMRMLALFV
]		1	j			AFAINFILLFYKVSTSSVVEGKELPTRSSSENA
]	}	j	Ì	ļ		KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL
1		.				PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK
1			ļ	İ		LEFDGLYITEQPSEDDIKGOWDRLVINTOSFP
1	ł	ł	ì	ł		NNYWDKFVKRKVMDKYGEFYGRDRISELLG
1 1	}	- 1	J		ļ	MDKAALDFSDAREKKKPKKDSSLSAVLNSID
			ì	l		VKYQMWKLGVVFTDNSFLYLAWYMTMSVL
		ì		ļ		GHY\NNFFFAAHLLDIAMGFKTLRTILSSVTH
	1	ľ	ì	1		NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF
	į	1	1		. }	YNKSEDGDTPDMKCDDMLTCYMFHMYVGV
	ļ	ļ	ŀ]	RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI
1	l	}	ľ		1	VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC
	ĺ		İ			FICGIGNDYFDTVPHGFETHTLQEHNLANYLF
	ľ	ł			ŀ	FLMYLINKDETEHTGQESYVWKMYQERCWE
1	ł	}	- 1			FFPAGDCFRKQYEDQLN
446	1796	A	3592	1	355	AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
""	1,70	^	2272	*	روو	DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL
1 1		l	ľ			SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML
		- 1	- 1		1	
<u> </u>						PRLVSNSWTQAILLPRPPKMLGLQV

					, <u> </u>	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide	i	in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	seq- uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	uence	1	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
ucia	1		""	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}	1	J	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ŀ				peptide	3 Aquana	/=possible micleotide deletion, \=possible
1			ĺ	sequence		nucleotide insertion
447	1797	A	3598	1202	1070	LFVGGGPICPEGASGFAPGPAPAPRVGVDAEV
1 '''	1	١	1 3370		1 .0.0	GR*V*GAAASQGA/GSLRPRPTGPGHPGAWL
1						QVWGAAAVCAGPAM*/AVRAKRGPRAG*EP
1	ļ	ļ	1		j	NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG
ļ			ļ		ļ	PSSRSAPGLASGPAAPLLQGVHSSAGPLLCYI
1	ļ		1			NGTLALGLKP**AWGWGEWRPKG
448	1798	Α	3604	3115	557	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE
i					1	GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG
1		ł				LLLVLLGIGFLVWHLQYRDVRVQKVFNGYM
1 .	}	l	ł		}	RITNENFVDAYENSNSTEFVSLASKVKDALKL
1 .	ł	t	ł		ļ	LYSGVPFLGPYHKESAVTAFSEGSVIAYYWSE
	}	l				FSIPQHLVEEAERVMAEERVVMLPPRARSLKS
;]	1]			FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR
	1					GVELMRFTTPGFPDSPYPAHARCQWALRGD
{		ļ	ł			ADSVLSLTFRSFDLASCDERGRHLV\TVYNT\L
1	{	İ	(SPMEPHA/LVQLCGTYPPSYNLTFHS/S/QNVL
1	ł	ļ				LITLITNTERRHPG\FEATFFQLPRMSSCGGRL
1						RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN
1	į	ļ	1	-		QHVKVRFKFFYLLEPGVPAGTCPKDYVEING
1	Ì]				EKYCGERSQFVVTSNSNKITVRFHSDQSYTDT
1			'			GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR
1			ĺ	ı		CDGWADCTDHSDELNCSCDAGHQFTCKNKF
1	1	l				CKPLFWVCDSLNDCGDNSDEQGCSCP\AQTF
1 .		i				RCSNGKCLSKSQQCNGKDDCGDGSDEASCP
} .		ł				KVNVVTCTKHTYRCLNGLCLSKGNPECDGK
						EDCSDGSDEKDCDCGLRSFTRQARVVGGTD
ļ		j	}			ADEGEWPWQVSLHALGQGHICGASLISPNWL
ľ						VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
			[]			QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL
1			'			ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV
						TGWGHTQYGGTGALILQKGEIRVINQTTCEN
] !			LLPQQTTPRMMCVGFLSGGVDSCQGDSGGPL
1			i .			SSVEADGRIFQAGVVSWGDGCAQRNKPGVY
						TRLPLFRDWIKENTGV
449	1799	A	3618	2	613	FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ
					,	EMTRRPSLMAGRQHGWSAQQSATVANPVPG
						ANPOLLPHFLGEPEDVYIVKNKPVLLVCKAV
					<u>'</u>	PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP
1						TMEVRINVSRQQVEKVFGLEEYWCQCVAWS
	,					SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL
450	1000		1200	- ,	0/2/	EQGIVLPCRPPEGIPPAE
450	1800	Α	3620	1	2676	MEPSLGQGMDLTCPFGVSPACGAQASWSIFG
] [İ		ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE
						PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA
}						TSPEGETDKNLANRVHSPHKRLSHRHLKVST
1 1						ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN
					İ	LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA
		ļ)			VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS
						SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP
					İ	RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA
						PVTNSSGKMALNSPQPGPVESELGKQLLKTG
1						WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP
						MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE
}						PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK
						APGLKDFQIQVQPVRMQKLTKLREEHILMRN
				į		QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP
						LTEKEVENVFVQLSSAFRNDSYTLESRINQAE
(ĺ	[1	RERNLTEENTEKELENFKASITSSASLWHHCE
.				İ		HRETYQKLLEDIAVLHRLAARLSSRAEVVGA
<u> </u>						THE TANKED IN THE TANKED THE TOTAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VRQEKRMSKATEVMMQYVENLKRTYEKDH AELMEFKKLANQNSSRSCGPSEDGVLRTARS MSLTLGKNMPRRRVSVAVVPKFNALNLPGQ TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE NGKTNGDPDCEASAPALTLSCLEELSQETKA RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ KSESPEEPEEVEETEEEEKDPRSSKLEELVHFL QVMYPKLCQHWQVIWMMAAVMLVLTVVL
	l.			!		GLYNSYNSCAEQADGPLGRSTCSAAQKDSW WSSGLQHEQPTEQ
451	1801	A	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI RHQRIHTGEKPYECKVCGKAFRVSSQLKQHQ RIHTGERPYQCKELKGRGAEMLAVLAVKEQ NRTPVNYGK
452	1802	A	3628	2	195	MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV QCTALGVWTAPAPVCIAVQCQHLEALNEGT MG*DYPFTAFAYGSSCKYECHTVYRVRGLD MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP VCIAVQCQHLEALNEGTMG
453	1803	A	3637	662	142	IQAKGLGIWHVPNKSPMQHWRKGSLLRYRT DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT DNGPVIPVVYDFGDAQKTASYYSPYGQREFT AGFVQFRVFNNERAANALCAGMRVTGCNTE HHCIGGGGYFPEASPQCGDFSGFDWSGYGT VHVGYSSSREITE\AAVLLFYR
454	1804	A	3641	1	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD AAGRDPGQDNDRNTAEPAFPPPPRVMAAAA ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	A	3646	2	414	AAAGRGASGALTGEGGGEQGRRVGLGSRAH SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE SRVRAPSYDDIT
456	1806	A	3656	396	8	QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK F
457		A	3660	14	1961	SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF TKCRSPERETFSCHWTDEVHHGTKNLGPIQLF YTRRNTQEWTQEWKECPDYVSAGENSCYFN SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ PDPPIALNWTLLNVSLTGIHADIQVRWEAPRN ADIQKGWMVLEYELQYKEVNETKWKMMDP ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIIF GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH EKLHINLGVKDGDSGRTSCCEPDILETDFNAH DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGQKNKAGMSQCDMHPEMVSLCQENF LMDNAYFCEADAKKCIPVAPHIKVESHIQP\S LNQEDIYITTESLT\TAAGSP\GTGEHVPGSEM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS
458	1808	A	3663	154	462	SCGYVSTDQLNKIMP TRAPASGRSGAGLALSANAPDSGGHPGATEG PAGSLAHASGSARGTWRVRGRGSHGWERTV GAGGCANPVPALHSCASAPRGTGRVSALGPK TGSSPLSSPKG
459	1809	A	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ SQNALGKYNTSMALFESNSFEKTILESPYYVD LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD FASPTYDLIKSGCSRDETCKVYPLFGHYGRF QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC VNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR SA\NGNSGFQHETHAEETPNQPFNSVHLFSFM VLALNVVTVATITVRHFVNQRADYQ\YQKLQ NY
460	1810	A	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA GLDLL\TSGDPPASTSQSARTTDVSHRAQPLAI S
461	1811	A	3671	2472	2099	IGVLAFETGSCSVTRLYCIGIMPHCSLDLAGS\ TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV VQTGL*LLALSNPPALASQIAGISGMSHRAWP GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	A	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQTN CYYD/STKSFFYISCG*K\RKPTWAENRRLNA KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG HGS
463	1813	A	3673	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV WPGQKPRPSQQQHQMCASPTLGQRSPFALEP VPAYHGGRDPFASARPSPVGIPKPRAAPAGG GWRRIRPKSSTK
464	1814	A	3676	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE APACRISFLPLTRLRRTESVPSDINNPVDRAAE PHFGTLPKALTKKEHPPAMNHILDSSNPSSTT FSTPSSPAPFPTSSNPSSATTPPNPSP\GQR\DSR FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE AADGTRLDDQPKADVLEAHEAEAEEPEAGK SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR WHGEVAIRLLEMDGHNQDHLKLFKKEVMN YRQTRHENVVLFMGACMNPPHLAIITSFCKG RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF \GISGVVP\EGRRENQLKLSHDWLCYLAPEIVR EMTPGKDEDQLPFSKAADVYAFGTVWYELQ ARDWPLKNQAAEASIWQIGSGEGMKRVLTS VSLGKEVSENLSACWAFDLQERPS\FSLLMD MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR FERFGLGVLESSNPKM
465	1815	A	3679	8	803	PERFOLGVLESSNIKM IPSPAWWNSTWADTFSLLLALAVALYLGYY WACVLQTHRAFCASNTEDLETVVNHIKHRYP QAPLLAVGISFGGILVLNHLAQARQAAGLVA ALTLSACWDSFETTRSLETPLNSLLFNQPLTA GLCQLVERLSY/E*DLQARTIRQFDERYTSVA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		Į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
[ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i		Ì		peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
Į .		1				FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA
1			ļ			DDPFSTVCALPKQAAQHSPYVALLITARGGHI
						GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE
466	1016	<u> </u>	3684	3	307	GLPDLRALLPSEDRNS SSOYTVOSKTKIFL*AAREKO/RHTCRRFSIRLS
466	1816	A	3084	3	307	ANISSOTGEARGOWPSVPKVLKEKKLSTKKS
1		}				FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
		l				TGVLQG
467	1817	A	3687	2465	837	ELPTPLIAAHOLYNYVADHASSYHMKPLRMA
407	1017	^	3007	2403	657	RPGGPEHNEYALVSAWHSSGSYLDSEGLRHO
Ī		[[DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF
						FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE
		[[APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG
						MARARLAQLVRLAGGHCRRDTLWKRLFLLE
	,	 				PPGPDRLRLGGRLALAELEELLEAVHAKSIGD
1		ł				IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR
1		1	:			HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH
1			·			LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY
		Ι.	ļ			HHLESVINTACFTLWTRLL*GSGLDH*MSLFL
		ĺ	İ :			ESWAYQIACQRQD*PALLGPRASQTLSDTKG
1		~				FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR
			·			ESGQPRGPLGPFWGTPFGPPGRVSGVHTGWQ
1	!		[İ	TPPRAPLPESCPL\PLTTVSHLCPLSLRVFTSHL
1			,			DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA
1		1	1		ì	LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG
1	1010		3601	0.00	400	GKQQRN
468	1818	A	3691	960	499	QTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI
		l	'			QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH
1		l				DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF PGNF*FLVKTGFPHVGQTGFELLTSSDLAPLA
						SONGGITGMSPCAWPFFFFFFGLC
469	1819	A	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP
1409	1017	^	3/14	4/4/	495	HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF
[[VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL
				*		SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG
1		}				SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL
						ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK
					,	RSGHVNIVEPSLMLLKGSLQPGMWESTWQK
					•	NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR
1]				ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK
						FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH
						FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS
						HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE
						DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY
1.						DFHLKYLLKTQENVYNIIEEVKKICSVLGCVE
1						TKQITDAVNELSLILQRKGENFYQSSETSAKG
						LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV
				,		PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR
						INFPLEIKSLPRESMLTVKLFGIACATNNANLL
						AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM
						ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD
						SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE
						EKKRYLWFYRFYCNNENCSLPLVLGSAPGW
1.						DERTVSEMHTILRRWTFSQPLEALGLLTSSFP
1			•			DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV
		·		• 1		KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL
]				.		LKNAENEAYFKSWYQKLLAALQFCAGKALN
1						DEFSKEQKLIKILGDIGERVKSASDHQRQEVL
						KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH DACSYFTSNALPLKITFINANLMGKNISIIFKA
L	L		L		L	DACS IF I SNALTLALITINANLMUANISHEKA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MITYRCLSTGKDQRLVQMVPDAVTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS EMEYFITEGG\KNPQHFQDFVELCCRAYNIIR KHSQLLL\NLL\EMMLYAG\LPELSGI\QDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL
470	1820	A	3718	430	75	TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPFTNSDHRRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM KSLEKDEFVGGMLLSNPIW SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA
						VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVOLLLSLGSVFLWALL
471	1821	A	3723	891	494	LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	A	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN *ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473	1823	A	3746	3	500	THASGGARSGAGWAGRGVRAGTEAGRGGIF LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824	A	3753		. 5262	RPLFAREGGIYAVLVCMQEYKTSVLVQQAG LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLLNLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSP'ERAALETPIIQGQDGSPELLIRSLV GGPSAELLLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL DQYLEQRETSRNPLSRAASFASRVRLCHLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
	ucilco	l	914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
uence	l	1	714	amino acid		T=Threonine, V=Valine, W=Tryptophan,
ſ	1	ĺ	(of peptide	
ļ		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
			L	sequence	L	nucleotide insertion
-	1					ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ
1	į]				QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG
			!			LELATTFEHFYQHYMADRLLSFGSSWLEGAV
	·					LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF
1	l					QLQRLDKLFLEQEDEEEKRL*EEEEEEEEA
1	l .					EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP
						RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG
1	ł		1			PHRRLQWTWLGRAELQFGKQILHVSTVQMW
	ł					
	1					LLLKFNQTEEVSVETLLKDSDLSPELLLQALV
1		•	1			PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ
1			1			RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL
1			İ			LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ
1	1					KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL
1						GQGYVKRRDDRPQILMYAAPEPMGPCRGQA
	1					DVPFCGSQSETSKPSPEAVATLASLQLPAGRT
1		}	 			MSPQEVEGLMKQTVRQVQETLNLEPDVAQH
	1	Ì	1			LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV
i			ĺ			HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC
1	ĺ		(MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD
		,	1			CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE
	Į.		j .			SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
1		[l
i	i					CGWASCFNCSFPEAHYPASCGHMSQWVDDG
	l					GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE
			i .			KNEGCLHMTCAKCNHGFCWRCLKSWKPNH
1						KDYYNCSAMVSKAARQEKRFQDYNERCTFH
Ī		,	1			HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA
						CQGLEQARKVLAYACVYSFYSQDAEYMDVV
1	!		1			EQQTENLELHTNALQILLEETLLRCRDLASSL
1	I					RLLRADCLSTGMELLRRIQERLLAILQHSAQD
	i					FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP
1	!		1			EAEEEEDDEDDVPEWQQDEFDEELDNDSFS
1	i					YDESENLDQETFFFGDEEEDEDEAYD
475	1825	Α	3754	1093	96	GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ
4'3	1023	Λ.	3134	1093	90	ATGRRRRTRTQQRTAALLTDGTTKTGAAW
	Į.		1			7.7
1	İ					SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN
	l .					PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP
1	1					DGTR\RPASITGVAQSPATRATPSLPCLHVPAP
1	1					SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA
	l ·					RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE
1	1					PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS
1				•		HVYIIRATINSISHPLCRAQSSPWEAAGVWRR
	İ					PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN
1						TLWEEGRORPPETLOPAR
476	1826	A	3758	901	521	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK
1			3.55			RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG
1	[FTMLARMVSIS*PRDLPALASQSAGITGVSHH
L	100=					APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG
1						ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI
	L		L		·	RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	Α	3763	267	1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP
						PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS
1						PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P
						GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY
		,				WTGHSFASQAWLRQVPEVSKHLQCPSAESLL
	,					TMEYHOPEDPAPGKAGTAEAVIPENHEVLAG
1						PDEHPQDTDARDADGEAREREP/RRPSFAA*P
				Ì	·	VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA
'				٠. ا		CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE
L						QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK DFFQKVSQVYVAIDERLASLKTDTFSKTREEK MEDIFAQKEMEEGEFKNWIEKMQARLMSSS VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS AMDASPRNISPGLQNGEKEDRFLTTLSSQSST SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG NSYNPIPFPFDPDKHYLMYYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE TLRGADSAYYQVGQTGKEGTENQGVEPQDE VDGGDTQKKQLINPHVELQFSDANAKFYCRL YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ ARGGKSGAAFYATEDDRFILKQMPRLEVQSF LDFAPHYFNYITNAVQQKRPTALAKILGVYRI GYKNSQNNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLLDENL LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD KKLEMVVKSTGILGGQG*MPTVVSPELYRTR FCEAMDNYFLMVPDHCTGLGLNC
480	1830	A	3777	251	3	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI KLNI\KDPAITLDVYPNEVKNYVRTKTYTQMF I/ANFIMAKSWKQPTHPSVRT
481	1831	A _.	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID SIEANAESSEVLVERAPGQLQRPA\YYQKKSR KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI VLPVSCFQGQKFN
482	1832	A	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S /L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAPSAAEKARRLWCS/VFNIERRNL/CEYPTK LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI QEEGKMF*KEKCFKRKE
484	1834	Α .	3798	1	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\ SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS PC*PGWS*SPDLVIRPPRLPKCWDYRREPPRP A*FFVFLVE\QGFTMLARMVSIS*PQ\CDLPAS VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG SWDYRRPPRPANP\CIFSRDGVSPC*PGWSRS PDLVIRPPRPKVLGLQA
485	1835	A	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF KQFSC\LSLPSS\WD*R\VPT\SRPA\KF\C\VIF*\DGV SHCQPG\WSAV\VQPPLH
486	1836	A	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771	320	FDPDWTRAAGIRHEKKPKALAYRRENSPGDL PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAPWSRP SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG RSRSRSQSRSQSRPGQKRREEPR

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FRACLLELIPYAPTLSWTACPPAMAGPRGLLP LCLLAPCLAGFSFVRGQVLFKGCDVKTTFVT HVPCTSCAAIKKQTCPSGWLRELPDQITQDCR YEVQLGGSMVSMSGCRRKCRKQVVQKACCP GYWGSRCHECPGGAETPCNGHGTCLLOGMDR NGTCVCQENFRGSACQECQDPNRFGPDCQSV CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD QELPVWQELGFPQNNPRLRKAPNCKCLPG*H RNGLIATPNPCRP
489	1839	A	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG QDGLDLL/NLMIHPPRPPKVLGFQA
490	1840	A	3825	79	9748	GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Ghttamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL
1						DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF
						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
	İ					KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
1						LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ
1						CHKENEDKWKRLSRQIADIILPMLAKQQMHI
1						DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED
						IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
1						EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
1						KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS
1						GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
1						ARSMITTHPALVLLWCQILLLVNHTDYRWW
						AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
						AKLGMCNREIVRRGALILFCDYVCONLHDSE
						HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS
						AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
			l j	j		HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL
						ACRRVEMLLAANLQSSMAQLPMEELNRIQEY
						LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
						PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
!		- 1				SQCWTRSDSALLEGAELVNRIPAEDMNAFM
						MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
		1				AREVTLARVSGTVQQLPAVHHVFQPELPAEP
1		1				AAYWSKLNDLFGDAALYQSLPTLARALAQY
		i				LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
				ļ		SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
						WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
		İ				EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI
					i	TAACEMVAEMVESLQSVLALGHKRNSGVPA
1		- 1				FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG
	•			-		WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
						INTLGWTSRTQFEETWATLLGVLVTQPLVME
		l	- 1			QEESPPEEDTERTQINVLAVQAITSLVLSAMT
		1	l			VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
				i		LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
						PVPSLSPATTGALISHEKLLLQINPERELGSMS YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE
	ĺ	ĺ	ĺ	Ī	<i>'</i> [ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL
		l				LELYSRWILPSSSARRTPAILISEVVRSLLVVS
			-			DLFTERNQFELMYVTLTELRRVHPSEDEILAQ
		1	i	1		YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
	1	1				RSSHLPSRVGALHGVLYVLECDLLDDTAKOL
1	ł	- 1		1	1	IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA
		1		İ		TAFYLIENYPLDVGPEFSASIIOMCGVMLSGS
		1		ĺ		EESTPSIIYHCALRGLERLLLSEQLSRLDAESL
]	į			j		VKLSVDRVNVHSPHRAMAALGLMLTCMYT
	. [Í	İ	ľ	GKEKVSPGRTSDPNPAAPDSESVIVAMERVS
				Į	}	VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ
			į	-	1	DIMNKVIGEFLSNQQPYPQFMATVVYKVFQT
				ŀ	į	LHSTGQSSMVRDWVMLSLSNFTORAPVAMA
]	1	1		İ	TWSLSCFFVSASTSPWVAAILPHVISRMGKLE
	ł	- 1	1	ł	1	QVDVNLFCLVATDFYRHQIEEELDRRAFQSV
		i			1	LEVVAAPGSPYHRLLTCLRNVHKVTTC
491	1841	A	3826	469	302	SNPPASASRVAGITGVHQHAWLIFVFLVEMEF
					<u> </u>	HHVGQAVLKLLISGDLPVSASQSA
492	1842	A	3836	392	88	VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\EE
1 1		- 1		1	.	FQSEWTAVV/P/EFTATQSEVADWFKDMQVP
			. 1			SVPIQQFPTEDWST*PTMNDWSATSTAQTTE
						WVRITTEWP
		_				

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	ľ	Į.		peptide	1	/=possible nucleotide deletion, \=possible
100	ļ	ــــــــــــــــــــــــــــــــــــــ		sequence		nucleotide insertion
493	1843	A	3838	19	380	TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
1	}	1		{	1	KFKRS*EKAHIRYKIDQPEDIPLK\EFLCKHSK
		1			}	CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL
104	1041	 	0035	<u> </u>	l	VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	A	3845	2	352	FFFLRRSL/DSVAQAEAQWL\ELGLLQAPPPGF
				ĺ	[KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG
						FTVLARMVLIS*PCDPPTLASQGTAITGMSYH
406	1945		2047	1554	<u> </u>	ARPQDIDFLYAHQGRCWFRLL
495	1845	A	3847	1774	40	DIFFRRAKEGMGQDEAQFSVEMPLTGKAYL
1						WADKYRPRKPRFFNRVHTGFEWNKYNQTHY
Í	1					DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL
ľ						EACADNKDFAILRFHAGPPYEDIAFKIVNREW
1						EYSHRHGFRCQFANGIFQLWFHFKRYRYRR*
1				!	}	RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL
1				'		QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC
ĺ						HGELRRHWDRLA*GPDATEGALGASFEHEG
						GQQPPADLTVQADTLHRPSARLGGAHRACPK RRPHRVLWRWARGAWAWRCQAREKQETQG
						QPCHITGHPLGREAEPAAAGAAPALAHRPPF
1	1				i	ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD
	,					WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN
ł						VMGTKSH*AVLPPPPSTGPGGQGLPEGWGLE
				•		KGEGLPPGIPPPGLLTGPW\SMRPVTPSFAHIR
-						TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR
			.			LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT
						SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK
					·	SFVLMELAYWQDRMFF
496	1846	A	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG
			ł	·		LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG
[LKLLTSSALPALASQSAEITGMSHRIWPLPLLR
						RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
497	1847	Α	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR
						LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS
						PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR
				·		RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ
100		1				VPIL
498	1848	A	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG
]						EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP
[]	· ']	j	J	ļ	.]	LPCLANF*FLVETGFHHVGQADLKLLTSGDP
100	1040		2062	400		PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	Α	3863	423	263	APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI
						KIGINLTKEVKYLYTENYITLMKEIK/DTDKW
}	}	ł	.			KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP
						MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S
		Ī]	İ		FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI
500	1050		2066		15046	LVGIALNL
500	1850	A	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP
		J]		DLRPWASDLDIMGDAEGEDEVQFLRTDDEV
	1			1	. 1	VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP
		- 1		l		TSNAQNVPPDLAICCFVLEQSLSVRALQEML
	ł				ļ	ANTVEAGVESSQGGGHRTLLYGHAILLRHAH
	İ	1	İ	- 1		SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE
			,	- 1	1	ACWWTMHPASKQRSEGEKVRVGDDIILVSVS
	1	- 1	1	ł	1	SERYLHLSTASGELQVDASFMQTLWNMNPIC
	1		İ	l	ļ	SRCEEGFVTGGHVLRLFHGHMDECLTISPADS
[ſ	1	1	[ĺ	DDQRRLVYYEGGAVCTHARSLWRLEPLRIS
		ŀ	l	l	į	WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR
l l			1	ı	i	I. V V V I I ANK A HIK A INH CHRISK HKI I IV ADVD
	j	- 1		1	ļ.	DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
├ ──				sequence		nucleotide insertion
i :						APDPKALRLGVLKKKAMLHQEGHMDDALSL
1						TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS
						GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV
						LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
					!	VNLLYELLASLIRGNRSNCALFSTNLDWLVS
						KLDRLEASSGILEVLYCVLIESPEVLNIIOENHI
						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV
						RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
ľ						IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
						ATHLRVGWALTEGYTPYPGAGEGWGGNGV
						GDDLYSYGFDGLHLWTGHVARPVTSPGQHL
1						LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
						NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF
						LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
						RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
						LERIREKLAENIHELWALTRIEQGWTYGPVRD
			1			DNKRLHPCLVDFHSLPEPERNYNLQMSGETL KTLLALGCHVGMADEKAEDNLKKTKLPKTY
						MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
		.				NGHNVWARDRVGQGWSYSAVQDIPARRNPR
1						LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
1						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
						QSGRWYFEFEAVTTGEMRVGWARPELRPDV
]		ļ		J		ELGADELAYVFNGHRGQRWHLGSEPFGRPW
1					·	QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
1						ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD
						VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS
[[KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
		Ī				LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
						NLRRSAGGWSEAENGKEGTAKEGAPGGTPQ
}		l				AGGEAQPARAENEKDATTEKNKKRGFLFKA
					į	KKVAMMTQPPATPTLPRLPHDVVPADNRDD
			1			PEIILNTTTYYYSVRVFAGQEPSCVWAGWVT
		- 1		j		PDYHQHDMSFDLSKVRVVTVTMGDEQGNV
[]		1				HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL
		l				VIGCLVDLATGLMTFTANGKESNTFFQVEPN
		l	1			TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA
						AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
1 1	i	1			ł	MPNHFLQVETRRAGERLGWAVQCQEPLTMM
		ł				ALHIPEENRCMDILELSERLDLQRFHSHTLRL YRAVCALGNNRVAHALCSHVDQAQLLHALE
		İ	1			DAHLPGPLRAGYYDLLISIHLESACRSRRSML
	1	}	l			SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP
	ļ			Ì		GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
[İ			•	ARLSPAIPLEALRDKALRMLGEAVRDGGOHA
				1		RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE
[ĺ	- [ĺ	ĺ	[DVKQILKMIEPEVFTEEEEEEDEEEGEEEDEE
		ļ	j	1	•]	EKEEDEETAQEKEDEEKEEEAAEGEKEEG
	1	ł	1	l	ļ	LEEGLLQMKLPESVKLQMCHLLEYFCDQELQ
[l	ľ	ł	Ī	ľ	HRVESLAAFAERYVDKLQANQRSRYGLLIKA
	ł					FSMTAAETARRTREFRSPPQEQINMLLQFKDG
	į.					TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD
	. [1	1	l	· · .	GEEEPEETTLGSRLMSLLEKVRLVKKKEEK
	İ			l	į	PEEERSAEESKPRSLQELVSHMVVRWAQEDF VQSPELVRAMFSLLHROYDGLGELLRALPRA
				1	.	YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
J }						ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
	1	ĺ	- 1	- 1	1	TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	vence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	0		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1 4040	,		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1						
1 1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 1				p e ptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
!						CYFCRISRQNQRSMFDHLSYLLENSGIGLGM
1						QGSTPLDVAAASVIDNNELALALQEQDLEKV
1						VSYLAGCGLQSCPMLVAKGYPDIGWKPCGG
1 1						ERYLDFLRFAVFVNGESVEENANVVVRLLIR
						KPECFGPALRGEGGSGLLAAIEEAIRISEDPAR
						DGPGIRRDRRREHFGEEPPEENRVHLGHAIMS
1 1						FYAALIDLLGRCAPEMHLIQAGKGEALRIRAI
1						LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPK
]					· ·	MSASFVPDHKASMVLFLDRVYGIENODFLLH
1				•		VLDVGFLPDMRAAASLDTATFSTTEMALAV
1						
1						NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS
1						MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR
						YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLL
1 1						TNHYERCWKYYCLPTGWANFGVTSEEELHL
1						TRKLFWGIFDSLAHKKYDPELYRMAMPCLC
						AIAGALPPDYVDASYSSKAEKKATVDAEGNF
						DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF
1		•				DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
}						KDKEIYRWPIKESLKAMIAWEWTIEKAREGE
1 j						EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL
1 [SAVTLSRELQAMAEQLAENYHNTWGRKKKQ
j l						ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
1						QELLKFLQMNGYAVTRGLKDMELDSSSIEKR
1 1	1					FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
1						EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
1 1	1					TPAKVLGSGGHASNKEKEMITSLFCKLAALV
1	1					RHRVSLFGTDAPAVVNCLHILARSLDARTVM
					•	KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
						· · · · · · · · · · · · · · · · · · ·
1						KVSQARTQVKGVGQNLTYTTVALLPVLTTLF
1 1						QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG
1	i					TTKNTYVEKLRPALGECLARLAAAMPVAFLE
						PQLNEYNACSVYTTKSPRERAILGLPNSVEEM
[CPDIPVLERLMADIGGLAESGARYTEMPHVIE
1 1			1			ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP
] ']			J	J		CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM
1 1						KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR
1			Į		ļ	AGKVVSEEEQLALEAKAEAQEGELLVRDEFS
i i						VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS
]		ļ				AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ
				ļ	ļ	NEINNMSFLTADNKSKMAKAGDIQSGGSDQE
1 I		l			ļ	RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN
] [·	Į		ļ		MCAPTDODLITLAKTRYALKDTDEEVREFLH
]			·i	i		NNLHLQGKVEGSPSLRWQMALYRGVPGREE
1 I			İ			DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS
[-	ļ			KKAVWHKLLSKQRRRAVVACFRMTPLYNLP
			ŀ			THRACNMFLESYKAAWILTEDHSFEDRMIDD
j. 1	. 1	ļ				LSKAGEOEEEEEEVEEKKPDPLHOLVLHFSRT
]. [· 1	-].			
1		ĺ	l		ļ	ALTEKSKLDEDYLYMAYADIMAKSCHLEEG
; l			1	ŀ		GENGEAEEEVEVSFEEKQMEKQRLLYQQARL
J 1			. 1		1	HTRGAAEMVLQMISACKGETGAMVSSTLKL
[•		ļ			GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
[]	1	- 1	ł			IQALMQTCSVLDLNAFERQNKAEGLGMVNE
	j	ł			1	DGTVINRQNGEKVMADDEFTQDLFRFLQLLC
1	1	ł				EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL
] [ŀ].	I	.	RLQESISDFYWYYSGKDVIEEQGKRNFSKAM
]		į		1	ľ	SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW
l J	ļ	Į	l	i	}	DAVVGFLHVFAHMMMKLAQDSSQIELLKEL
		ſ		. [ĺ	LDLQKDMVVMLLSLLEGNVVNGMIARQMV
				ļ	ĺ	DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
1				t t		
		ı	ŀ	[l	QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI

<u> </u>	T OF A TO	1 17.	T 050	1 80 . 40		1 4 10 10 10 10 10 10 10 10 10 10 10 10 10
SEQ ID	SEQ ID NO: of	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	peptide	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghutamic Acid,
eotide	seq-	1	in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uc	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j	!	J	ļ	peptide	Jaquana	/=possible nucleotide deletion, \=possible
	ŀ	1		sequence		nucleotide insertion
	 	 		Joques		QFLLSCSEADENEMINCEEFANRFQEPARDIG
		1	l			FNVAVLLTNLSEHVPHDPRLHNFLELAESILE
	ł	1			Į	YFRPYLGRIEIMGASRRIERIYFEISETNRAQW
ļ	ļ	ļ	į			EMPQVKESKRQFIFDVVNEGGEAEKMELFVS
ļ .	l	ł				FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA
1		i				AEAGAEGAEGAAGLEGTAATAAAGATARV
						VAAAGRALRGLSYRSLRRRVRRLRRLTAREA
İ	}					ATAVAALLWAAVTRAGAAGAGAAAGALGL
1		ł	1	1	}	LWGSLFGGGLVEGAKKVTVTELLAGMPDPT
			}			SDEVHGEQPAGPGGDADGEGASEGAGDAAE
						GAGDEEEAVHEAGPGGADGAVAVTDGGPFR
	1				}	PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG
		İ			}	VDGVEEELPPEPEPEPEPELEPEKADAENGEK
	İ	ĺ		ĺ		EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG
						ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN
		ļ		İ		FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS
į			1			GGSSGWGLGAGEEAEGDEDENMVYYFLEES
İ						TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP
						LVIFKREKELARKLEFDGLYITEQPEDDDVKG
1		l			}	QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG DIYGRERIAELLGMDLATLEITAHNERKPNPP
						PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG
						WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL
						RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV
						AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL
	'	ĺ	1			FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV
						FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV
					;	KEDMETKCFICGIGSDYFDTTPHGFETHTLEE
						HNLANYMFFLMYLINKDETEHTGQESYVWK
						MYQERCWDFFPAGDCFRKQYEDQLS
501	1851	Α	3869	467	665	VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK
		1			,	LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV
						N
502	1852	Α	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD
ļ				·		YRHAP\PLLTNF*FLVEMGFCYVGQAGRKLL
		1				ASSDQSALASQSAGITGISTAPGPPFFFLNFEA
						GSCSVAQAGVQ
503	1853	A	3891	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR
1	i					QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL
					!	VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP
]						HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET
	<u> </u>					WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC
						APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL
504	1054	- B	2004	270	30	L
304	1854	В	3896	279	70	MVSKSKSILMSYNHVELTFSDMKKMPEAFRR
}]				TQKHTTYLIPYQVIFWSTGKDAMRSFMMPFY
505	1855	_	3899		1206	QKEYYENQ*
دەد	1033	A	ללסכ	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG
						NENTKLELRKVPPELNNISKLNEHFSRFGTLV
						NLQVAYNGDPEGALIQFATYEEAKKAISSTEA
						VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL
				l		VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS
*	-					ASSDLPQVLST\LLA*QKQCIIQLL/WKAAQKT LLVSTSAVDNNEAQKKKQEALKLQODVRKR
						KQEILEKHIETQKMLISKLEKNKTMKSEDKAE
				l		IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI
						KTKTQMQKELLDTELDLYKKMQAGEEVTEL
					ļ	RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA
						VHGRGRGRGRGRGVPGHAVVDHRPRALEIS
		J				TAGEORGICONORO TA OHAT TOHIC INDEED

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methiomine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Trytophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA VITFKTRAEAEAAAVHGARFKGQDLKLAWN
506	1856	A	3911	1952	919	KPVTNISAVETEEVEPDEEEQREIIIA DAELSGTLSLVLTQCCKRIKDTVQKLASDHK DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA DSQRLLNEVMVEHFFRQGMLDVAEELCQES GLSVDPSQKEPFVELNRILEALKVRVLRPALE WAVSNREMLIAQNSSLEFKLHRLYFISLLMG GTTNQREALQYAKNFQPFALNHQKDIQVLM GSLVYLRQGIENSPYVHLLDANQWADICDIFT RDACALLGLSVESPLSVSFSAGCVALPALINIK AVIEQRQCTGVWNQKDELPIEVDLG*KSAGY HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	A	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA PGSGARCHPPSTCSPSWASPG*GAKASPALPR SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ RPAKQRDKRNRHLGR
508	1858	A	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT TRHGASRGGVCAHSIEGGELFERIVDEDYHLT EV
509	1859	A	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	A	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT SASSVAATTG
511	1861	A	3956	1	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV RWQPSEKQPPPPAHRGPADSLSTAAGAAELS AEGAGKSRGSGEQDWVNRPKTVRDTLLALH QHGHSGPFESKFKKEPALTAVARTARKRKPS PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRM\P\SP MAALILVADNAGGSHASKDANQVHSTTRRN SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL EPVHPASLPDSSLATSAPLCCTLCHERLEDTH FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC PSGEKCPLVGSNVPWAFMQGEIATILAGDVK VKKERDS
512	1862	A	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR VAGTTDTHHHTWLILGSSVQTGFDHVGQAG LELLTSGDPPISASESAGIMGMSHCVWP*SWG LSHHMAPPQGDGGRARGTPGPEQSFWNLSC H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPLFHHGL*AWLWCPELLLQGQARRH*RSPPS/FK CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H PPICHWPSRRSLGDPLLPRSQG*RDGT*ASTFC SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP PSRPDRSRNSNSLSR
513	1863	A .	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTTDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

nucle eotide sequence	RCG DSSI KHN HRT LQQC ELKR LERT OMAG DLGD
### A ### A	RCG DSSI KHN HRT LQQC ELKR LERT OMAG DLGD
SLASSTVGLAGQVVHTETTEVVLTADP GIQLQGSVFATETLSSPLISYHEADSPAE VLQiGDRVMANNGIPTEDSTFEEASQLIJ TSKVTLEIEFDVAESVIPSSGTFHVKLPK VELGITISSPSSRKPGDPLVISDIKKGSVA GTLELGDKLLAIDNRLDNCSMEDAVQ EDLVKLKIRKDEDNSDEQESSGAITYTVI YGGPLGITISGTEEPPFU**IISSLTKGGL GAIHIGDRILAINSSSLKGRLSEAIHLLI ETVTLKIKKQTDAQSASSPKKPFISSHLS VEEDSSPAQKPGKLSDMYPSHGCPSVD SWDGSAUDTSTYGTEGTSFQASGYVNFN WRSPKQRGSU.SPVTNKPRSQTYPDVGLS WDRSTASGFAGAAVDSAETQEENFWS DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSSRPHYSQTTRSD DVGRKSVTLRKMKQEIKEIMSPTPVELF LYKDSDMEDFGFSVADGLLEKGVYVK\ GPGDLGGLKPYDRLLQVNHVRTRDFD VPLIAESGNKLDLVISRNPLASQKSIDQC D**SEQNSAFTQQPSHGGNLETREPTINTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIEFFAVLINTSTGQIDSEFQAYVQP LSFCMELTGIKQAQVDEGVPLKICLSQ WHKIQQQKNIII**ATGISEPS/DF*SKIMC VR*RISYTY**SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG***EGD DLLGPEFLARRLFKTKTVGST**SISKN PNFIIEEGTDLIF**QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLCQKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDULGAEDNEWVYRLY DKDNLYFVANGFIHRDIKPD	RCG DSSI KHN HRT LQQC ELKR LERT OMAG DLGD
GIQLQGSVFATETLSSPPLISYTEADSPAE VLQIGDRVMANIGPTEDSTFEASQLLI TSKVTLEIEFDVAESVIPSSGTHVKLJFK VELGITISSPSSRKPGDPLVISDIKKGSVA GTLELGDKLLAIDNIRLDNCSMEDAVQ) EDLVKLKRKDEDNSDEQESSGAITYTV YGGPLGITISGTEEPFDL*IISSLTKGGL GAHIGDRRLAINSSLKGKPLSEAIHILL ETVTLKIKKQTDAQSASSPKKPPISSHLS VEEDSSPAQKPGKLSDMYPSHGCPSVD SWDGSAIDDTSYGTEGTISFQASGYVNFN WRSPKQRGSLSPVTKKPRSQTYPDVGLS WDRSTASGFAGAAIDSAETEQEENFWS DLETCCQSGILRELAITMSGSTMSLINH PRSPAGSBRPSFQERSSRPHYSQTTRSN DVGRKSVTLRKMKQEIKEIMSPITVELE LYKDSDMEDFGFSVADGLLEEGVYVK GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGINKLDLVISRNPLASQKSIDQC O*SEQNSAFFQQPSHGGNLETREPTIVIL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSLAI NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLINTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WHKIQQQKNIIFATGISEPSDF*SKIMC VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGD DLGPEFLARKPLFKKTYQSTF*SISKN PNFIIEEGTDLIP*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYPVMDYIPGGDMMSLLRIMGIF ARFYLAELTCAVESVHKMGFIHRDIKPD	RCG DSSI KHN HRT LQQC ELKR LERT OMAG DLGD
TSKVTLEIEFDVAESVIPSSGTFHVKLPK VELGITISSPSSRKPGDPLVISDIKKGSVA GTLEIGDKLLAIDNIRLDNCSMEDAVQI EDLVKLKIRKDEDNSDEQESSGAIIYTVI YGGGPLGNTISGTEEPFDL*IBSLTKGGLL GAHIGDRILAAINSSLK GKPLSEAHILL ETVTLKIKKQTDAQSASSPKKFPISSHLS VEEDSSPAQKPGKLSDMYPSHGCPSVD SWDGSA\IDTS\YGTEGTISFQASGY\NFN WRSFKQGSULSPYTKPRSQTYPDVGLS WDRSTASGFAGAAIDSAETEQEENFWS DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSSRPHYSQTTRSN DVGRKSVTLRKMKQEIKEIMSPTPVELL LYKDSDMEDFGFSVADGLLEKGVYVX GFODLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTIL 514 1864 A 3967 833 800 LEKQCVSGMATKKILARQLGLIRRKSIAI NLGRSKSKQLFDYLLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WHKIQQQKNIIFATGISEPS/DF*SKIMC VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI DLLGPETLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIF*QVKHNPCHRLTPEGG NRADS 516 1866 A 3977 2 1357 KMC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERADILAEADNEWVYRLI DKDNLYYMDYYPGGDMMSLLIRMGIF ARFYLAELTCAVESVHKMGFIHRDIKPD	KHN HRT LQQC ELKR LERT (MAG DLGD
VELGITISSPSSRKPGDPLVISDIKKGSVA GTLELGDKILLADNRILDNCSMEDAVOJ EDLVKLKIRKDEDNSDQESSGAIITYV YGGPLGVITISGTEEPFDL*IBSSLTKGGLL GAHIGDRILAANSSLKGKPLSEAHHLL ETVTLKIKKQTDAQSASSPKKFPISSHLS VEEDSSPAQKPGKLSDMYPSHGCPSVD SWDGSA\IDTS\YGTEGTISFQASGY\NFN WRSFKQRGSULSPYTKFRSQTYPDVGLS WDRSTASGFAGAADSAETTGEERFWS DLETCQQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSRPHYSQTTRSN DVGRKSVTLRKMKQEIKEIMSPTPVELF LYKDSDMEDFGFSVADGLLEKGVYVKI GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNFLASQKSIDQC D*SEQNSAFFQQPSHGGNLETRETNTIL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAN NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WHKIQQQKNIFATGISEPSDF*SKIMC VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGD DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEECTDLIF*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAEDILAEADNEWVYRLY DKDNLYFVMDYPGGDMMSLLIRMGIF ARFYLAELTCAVESVHKMGFIHRDIKPD	HRT LQQC LLKR LERT (MAG DLGD
GTLELGDKLLAIDNIRLDNCSMEDAVQI EDLYKLKIRKDEDNSDQESSGAIIYTVI YGGPLGVITISGTEEPFDL*IISSLTKGGL, GAHIGDRILVAINSSSLKGKPLSEAIHLL ETVTLKIKKQTDAQSASSPKKPISSHLS VEEDSSPAQKPGKLSDMYPSHGCPSVD: SWDGSAIDTSVIGTEGTSFQASGYNFN WRSPKQRGS\LSPVTKPRSQTYPDVGLS WDRSTASGFAGAADSAAFTEQEENFWS DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSSRPHYSQTTRSS DVGRKSVTLRKMKQEIKEIMSPTPVELI- LYKDSDMEDFGFSVADGLLEKGVYVKI GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEILEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS/DF*SKIMC: VR*NISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGD] ULGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFY*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYLAELTCAVESVHKMGFIHRDIKPD	LQQC LKR LERT (MAG DLGD
YGGPLGVITISGTEEPFDL*IISSLTKGGL. GAIHIGDRILAINSSSLKGKPLSEAIHLL ETVTLKIKKQTDAQSASSPKKFPISSHLS VEEDSSPAQKPGKLSDMYPSHCCPSVD: SWDGSA\IDTS\YGTEGT\SFQASGY\NF\ WRSPKQRGS\LSPVT\KPKSQTYPDVGL; WDRSTASGFAGAA\DSAETEQEENFWS: DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSRPHYSQTTRS\\ DVGRKSVTLRKMKQEIKEIMSFTPVELL; LYKDSDMEDFGFSVADGLLEKGVYVK; GPGDLGGLRYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQVSGMATKRLARQLGLIRRKSLAH NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP- LSEFCMELTGIKQAQVDEGVPLKICLSQ WHKKIQQQKNIIFATGISEPS/DF*SKIMC: VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI DLLGPEFLARKPLFKTTYQSTF*SISKN PNFIIEEGTDLIFI*QVKHNPCHRLTPEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYLAELTCAVESVHKMGFIHRDIKPD	LERT (MAG DLGD
GAIHIGDRILAINSSSLKGKPLSEAIHLLG ETVTLKIKKQTDAQSASSPKKFPISSHLS VEEDSSPAQKPGKLSDMYPSHGCPSVD: SWDGSAIDTSYGTEGTSFQASGYNFN WRSPKQRGS\LSPVT\KPRSQTYPDVGLS WDRSTASGFAGAADSAETEQEENFWS DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSRPHYSQTTRSN DVGRKSVTLRKMKQEIKEIMSFTPVELL LYKDSDMEDFGFSVADGLLEKGVYVKI GPGDLGGLKPYDRLLQVNHVRTRDFDO VPLIAESGNKLDLVISRNPLASQKSIDQQ D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEILEFPAVLLNTTSTQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WHKIQQQKNIIFATGISEPS/DF*SKIMC VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFI*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNILYFVMDYIPGGDMMSLLIRMGIF ARFYLAELTCAVESVHKMGFIHRDIKPD	MAG DLGD
VEEDSSPAQKPGKLSDMYPSHGCPSVD: SWDGSAUDTSYGTEGT\SPQASGY\NFN WRSPKQRGS\LSPVT\KPRSQTYPDVGLS WDRSTASGFAGAA\DSAETEQEENFWS DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSSRPHYSQTTRSN DVGRKSVTLRKMKQEIKEIMSPTPVELL LYKDDMEDFGFSVADGLLEKGVYVKI GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI NLGRSKSQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQDDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS\DF*SKIMC VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIF*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYLAELTCAVESVHKMGFIHRDIKPD	
SWDGSA\IDTS\YGTEGT\SFQASGY\NFN WRSPKQRGS\LSPYT\RPRSQTYPDVGLS WDRSTASGFAGAA\DSAFTEQEENFWS\ DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSSRPHYSQTTRS\ DVGRKSVTLRKMKQEIKEIMSPTPVELF LYKDSDMEDFGFSVADGLLEKGVYVK\ GPGDLGGLKPYDRLLQVNHVRTRDFDC\ VPLIAESGNKLDLVISRNPLASQKSIDQ\ D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI\ NLGRSKSKQLFDYLLVIDFESTCWNDGK\ SQEIIEFPAVLLNTSTGQIDSEFQAYVQP\ LSEFCMELTGIKQAQVDEGVPLKICLSQ\ WIHKIQQQKNIIFATGISEPS\DF*SKIMC\ VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI\ DLLGPEFLARKPLFKTKTYQSTF*SISKN\ PNFIIEEGTDLIF*QVKHNPCHRLTPEG\ NRADS 516 1866 A 3977 2 1357 KMLCQKESNYIRLKRAKMDKSMFVKII\ GAFGEVCLARKVDTKALYATKTLRKKI\ RNQVAHVKAERDILAEADNEWVVRLY\ DKDNLYFVMDYIPGGDMMSLLIRMGIF\ ARFYLAELTCAVESVHKMGFIHRDIKPD	
WRSPKQRGS\LSP\T\KPRSQTYPD\GLS WDRSTASGFAGAA\DSA&TEQEENFWS\ DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFQERSSSRPHYSQTTRSN DVGRKS\TLRKMKQEIKEIMSPTP\VELI LYKDSDMEDFGFS\VADGLLEKGVY\K\ GPGDLGGLKP\DRLLQ\VNH\VRTRDFDQ\ VPLIAESGNKLDL\VISRNPLASQKSID\QQ\ D*SEQNSAFF\QQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEK\QVS\GMATKRLAR\QLGLIRRK\SIAI\ NLGRSKSK\QLFD\YLI\VIDFESTC\WNDGK\ SQEIIEFPA\VLL\NTSTG\QIDSEF\QA\Y\QP\ LSEFCMELTGIK\QA\QVDE\GVPLKICLS\Q\ WHKI\QQ\KNIIFATGISEPS\DF*SKIMC\ VR*RISYTY*SKHKSK\GC 515 1865 A 3969 492 182 CR\F\WGISTHCDTCDPLSP\QTFG*E\GD\ DLL\GPEFLAR\KPL\FKTKT\QST\F\SIS\K\ P\NFIIEE\GTDL\IF*\QV\KH\NPCHRLTPEG\ NRADS 516 1866 A 3977 2 1357 KMLC\QKES\YIRL\KR\KMD\KS\MF\VKI\ GA\F\GU\CV\KY\T\KAL\Y\AT\KT\LR\KI\ R\QV\AH\VK\AE\GI\LAE\AD\NE\WV\RL\Y\L\ DK\DNL\YF\WD\YIP\G\GD\MSLLIR\MGIF\ AR\F\IAELTCA\VE\YH\KM\GFI\H\DIK\PD\	
DLETCGQSGILRELEATIMSGSTMSLNH PRSPAGSDRPSFOERSSSRPHYSQTTRSN DVGRKSVTLRKMKQEIKEIMSPTPVELF LYKDSDMEDFGFSVADGLLEKGVYVVLF GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTL S14 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAF NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WHKIQQQKNIIFATGISEPS/DF*SKIMC VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFN*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
PRSPAGSDRPSFQERSSSRPHYSQTTRSN DVGRKSVTLRKMKQEIKEIMSPTPVELF LYKDSDMEDFGFSVADGLLEKGVYVKI GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAF NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS/DF*SKIMC VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDD DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFV*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
DVGRKSVTLRKMKQEIKEIMSPTPVELI- LYKDSDMEDFGFSVADGLLEKGVYVKI GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP- LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS/DF*SKIMC: VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGD: DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFV*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
GPGDLGGLKPYDRLLQVNHVRTRDFDC VPLIAESGNKLDLVISRNPLASQKSIDQC D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS/DF*SKIMC: VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGD: DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFV*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	KVT
VPLIAESGNKLDLVISRNPLASQKSIDQQ D*SEQNSAFFQQPSHGGNLETREPTNTL 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAI NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS/DF*SKIMC: VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGD: DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFV*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
S14	
NLGRSKSKQLFDYLIVIDFESTCWNDGK SQEIIEFPAVLLNTSTGQIDSEFQAYVQP LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS/DF*SKIMC: VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGD: DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIF*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	ANIO
SQEIIEFPAVLLNTSTGQIDSEFQAYVQP- LSEFCMELTGIKQAQVDEGVPLKICLSQ WIHKIQQQKNIIFATGISEPS/DF*SKIMC: VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIF*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
WIHKIQQQKNIIFATGISEPS/DF*SKIMCT VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDI DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFV*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	QEHPI
VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDDLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIFV*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
DLLGPEFLARKPLFKTKTYQSTF*SISKN PNFIIEEGTDLIF*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKII GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	Ç1L
PNFIIEEGTDLIF*QVKHNPCHRLTPEEG NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKI GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
GAFGEVCLARKVDTKALYATKTLRKKI RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
RNQVAHVKAERDILAEADNEWVVRLY DKDNLYFVMDYIPGGDMMSLLIRMGIF ARFYIAELTCAVESVHKMGFIHRDIKPD	
ARFYIAELTCAVESVHKMGFIHRDIKPD	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
RQDSMDFSNEWGDPSSCRCGDRLKPLE	
RQHQRCLAHSLVGTPNYIAPEVLLRTGY CDWWSVGVILFEMLVGQPPFLAQTPLE	-
KVINWQTSLHIPPQAKLSPEASDLIIKLC	RGPE
DRLGKNGADEIKAHPIF*NQFDFSQ*PEI	SRS
AFKQFP*NHTTPTDTSNFDP\VDPDKLW EEENVNDTLNGWYKNGKHPEHAFYEFI	
FDDNGYPYNYPKPIEYEYINSQGSEQQS	
QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFV	RCV
1867 A 3980 1338 1022 FFFRRFTQSLGFLLFSFSFLFSCFFFFFFFV VFLDRVPLCHPGWSAVVQSQVT/VNLPI	
*RCRPPH/LANLCNFCRD\SFTTLPRLVLY	
QAIFQPQPPKVLGLQV 518	
SCFSLPE*LGYRHVPPCLANSVFSVEMG	K*E
VGQAGLELLTSGDLPALASQSAGITG\SF	
PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSI	FLH
LISINQGHNALWKAAG\PLPLKAGYC\QS	FLH IRAR
DSLKYG\SWDEKDLTVPQRDTHKRSVLI QRGK\LAVEMEEGHCLL\LPLGTECLGIK	FLH IRAR IHHS FSPC
HLFSSEMGE\NRPMVG\ARHVYSNAALL	FLH IRAR IHHS FSPC WIS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	aa		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ł	J	7.14	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
I				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	5045-000	/=possible nucleotide deletion, \=possible
İ		l		sequence		nucleotide insertion
 		·		•		LRCLGGEKHKSGLHARPVIVPSLELHYDMDSI
ļ]	AHV\FADLLLITLPSYYIPFC
520	1870	Α	3999	882	698	OSFRLSLLSSWDYRHM*PRLANF*TVFFCRDR/
						SLALLPRLVSNSWPQAILPPRPPKVLGLQT
521	1871	A	4011	1346	1178	FFF*ETVSCSAS*AGVRSHDNSSLQPPSPG\SSN
1		1			ļ	PPTSASHVAGATGTHHHAWLLSV
522	1872	Α	4015	2	377	QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
		Ì]	IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR
		l .				EYGPVYSTWSALEGELAEPLEGVSACIGNCST
						AL*ELTDDMTEDFLFVLREYILYSDSMK
523	1873	Α	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP
		1				KVKEVCKTSKS/GQVIYKGVSIRLRANFLAEP
		1				L*NRREWDEAIKVLKEKQ\FLSKMVYPANLSF
						GNEGDITSFPAK
524	1874	Α .	4020	1067	743	FFLRWSL/DSVAQAGVKWCNLGSLQAPPPGF
						TPFSCLSLPSSWDYRHPPPRLAN*LTNFLCF**
						RQGFTVLARMVLIS*PHDLPASASQSAGITGL
505	1875		4001	701	261	SHCSWPTSSILS
525	18/5	Α	4021	781	351	QFRVIFFFLRRSHSVAQAGMQWHDHSLLQPL
1		ļ				PPRLKQ/F/SHLSPPSIWDYRRVPPCLVNFSIFF
						VETGSCQPCLQLLGSSNPPASASQSAGIAGISH
1			i			QGQPE*SFDIRFACVIAALRETFQCLCSASRVN
526	1876	Α	4024	80	341	NKIINRPTHPVESSF TPSSTSRGTEEQOSSKMAWORREEKEHLNVR
320	1070	Λ	7024	80	341	RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFO
1						LHIHSSESQLHHSVKSPPSLSFRLM
527	1877	A	4026	593	230	DFYLYPERKKRGOMMTAVSLTTRPOESVAFE
1			.020		250	DVAVYFTTKEWAIMG\PAERALYRDVMLEN
						YGGCGPL*CHPTSKPALVFS\LEQGKESCFSPA
						TGSSLSRNDWRAGWIGYLELRRYTYLS
528	1878	A	4028	1160	242	GTSELLCIQRWNWGPAFPPRPGLALAPTLQLL
						VEMGSAKSVPVTPARPPPHNKHLARVADPRS
1 1						PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAQ
						DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE
						VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT
			•			QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT
						ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\
1 1				ĺ		NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA
						AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA
L	1050					WEQGQD\HDKENQHFPLVES
529	1879	Α	4039	2	366	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK
				-		CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T
i i			- 1	Ì	' i	ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV
520	1000		4055	250		HTEICT*MFIAVLFVVVKTWKQF
530	1880	A	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL
]			j	i	1	RKYGSRINLLKSKHDKNICTENYKT*MKEIEA
1	ſ	1	ĺ	ſ	ſ	/DTDKWKDILCSWIRRIHMKDILCSWIGRTHV
531	1001		4061	50	270	VKISILPKVNYRFYLISIKIIMAI
1221	1881	Α	4061	50	278	TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY
						HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH
		j				IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR
532	1882	A	4069	19	368	
332	1002	^	4009	17	200	NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY
ļ l			i	1		
	ŀ					KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT
533	1883	A	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKU
	1000	^	70/0	. 1	ا	ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR
				[KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC
						THE THE PROPERTY AND AND AND AND AND AND AND AND AND AND

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
534	1884	A	4088	sequence 3	1931	INAIPIKMPMMCMAKIEKNSS IIDSSTRRMESERSPLYRQLIDLGYLSSSHWNC GAPGQDTKAQSMLVEQSEKLRHLSTFSHQVL QTRLVDAAKALNLVHCHCLDIFINQAFDMQR DLQITPKRLEYTRKKENELYESLMNIANRKQE EMKDMIVETLNTMKEELLDDATNMEFKDVI VPENGEPVGTREIKCCIRQIQELIISRLNQAVA NKLISSVDYLRESFVGTLERCLQSLEKSQDVS VHITSNYLKQILNAAYHVEVTFHSGSSVTRM LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI ESLSASKLAKSICSQFRTRLNSSHEAFAASLRQ LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS LESRSLQDVLLHRKPKLGQELGRGQYGVVYL CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA VLLIMERLHRDLYTGLKAGLTLETRLQIALDV VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK YDNSVDVYAFGILFWYICSGSVKLPEAFERCA SKDHLWNNVRRGARPERLPVFDEECWQLME
535	1885	A	4090	2	417	ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS\ NSEQPNRGLDDST ALMPHEANYEEIFLKTDKDMDGFESGLEVRE IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD HFALAFHLIT\QKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK
536	1886	A	4102	569	829	HNRKRIWLRA DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV
537	1887	A	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE
538	1888	Α	4109	141	314	IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMNAIGLIITALPVS
539	1889	Α	4111	268	1	ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL
540	1890	A	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYTYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG
541	1891	Α	4146	282	778	SGAIVVQLAQPYMIGSIRVLLWDCDDRSY GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC

COEO TO	I OFO TO	1 1/-4	LOTA	1 Your 1 '-4 - 4	N-3'-1-3	I A Second Secon
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		l in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ļ	l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}	1			residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
ł	İ				sequence	
		İ	ł	peptide	1	/=possible nucleotide deletion, \-possible
			1	sequence	l	nucleotide insertion
			1		,	HAESENFAFWQDMKWKNKFWGKSLEIVPVG
1			1			TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW
1	1					IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN
1	ł		J			
1	1		1			VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG
<u></u>		 				PTPGGQCIWKP
542	1892	A	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA
	i					QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR
	l					INHLIFRGGAQITFLATFDDSPKAVLGDRLLLT
			i			ANVSSENNTPRTSKTTFQLELSVKDAVYTVV
}	l]	SSH
543	1893	A	4153	670	12	L
243	1073	^	4133	678	11	TISYPQCLTQMYFLISFANVDTFLLPIMALDH
I		i			1	YVAICSALQ*CSITTP/ELCQGLPVLA*AGSSLIS
I		ŀ				PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA
1		l				CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA
1		l	1			AAILRIPSPTRRRKACSICSSHLSLVTLFYGTV
		1				LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFTY
						SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	Ā	4158	3	538	
344	1894] A	4138	3	238	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP
j		[SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL
1	<u>}</u>	1	i			LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS
1		l				LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG
•		ł	1			LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE
]	j .					LKVGREGHVLPWQAHVVEF
545	1895	A	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE
343	1033	^	4100		412	
		1				DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM
		1				QDTASAMPCLPYYPTSHCFMAGGKSRSQGW
		J				ELELSGEPAPGWQVLAGYTYTQARYLRDASE
			1			ANVGQPLRPVDPR
546	1896	Α	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL
						QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV
		ł		-		FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA
		l				
	1005			2020		GITGVSHHA*PRASGRRCW
547	1897	A	4176	3029	1	AGPDGLAAPASCQGARGQTRVPGAFSWLAP
					1	GSHHASEGLAPGVPPAGGVSAQELTAPPQEG
,						WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA
1						RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS
1			[GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ
						PSRLPRLTPAOLLOGPGVLAAPPGP*HVPGFL
]						AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP
j i						HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS
1						TMAPIPSALAVWEPAGSSPQLSSAPADSS\PLP
						ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/
			1			RCPPACSPAAASSFSFESQPCPSAPSKASPAPA
						AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL
] .]						AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA
1 : 1						LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/
1						
]				į		PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP
į l						AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ
				ļ		ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP
1				1		PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP
					ŀ	RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC
				l		LOPLHLRAAOPLDPCCSLSPPGPPLPVPALPS
				l		WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL
] !				l	ł	PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ
				ļ	•	
]					}	PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL
j ł				ļ	j	TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP
1				J	ļ	MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP
j l				1	ļ	P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT
} [[ſ		PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA PGTWQPWPVSPAPLSCLSAWDPWELPSPQPQ VCSTAELPTSCLLSSPGPPAFQPPRFGCL*GPP
548	1898	A	4180	2369	844	GPPGLPPLQSSLSFPPPPPPPPPPPPPPAPALQWG LHLPGGRTK RIHREEDFOFILKGIARLLSNPLLQTYLPNSTK KIQFHQELLVLFWKLCDFNKVGQPRGALQGD GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR ADQSRVGLMHIGVFILLLLSGECNFGVRLNKP YSIRVPMDIPVFTGTHADLLIVVVFHKIITSGHQ RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP TIHKALQRRRTPEPLSRTGSQGGAPPWRAPA PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR MAPCGPWNLSPSRAWRMAARLRGSPARHGG SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ HGTLVGLLPVPHPILIRKYQANSGTAMWFRT YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	A	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE LLVRKWRVKSALGAMGQWQLEVGDPAPLG AGNLGPELIKESNANPIFMRKDTKMSFQWRIR NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	1900	A .	4192	1	1980	RHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL GASAMRRSEVLAEESIVCLQKALNHLREIWE LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE SLKERLIKSISVCQKELNTLCSELHVEPFQEEG ETTILQLEKDLRTQVELMRKQKKERKQELKL LQEQDQELCEILCMPHYDIDSASVPSLEELNQ FRQHVTTLRETKASRREEF/VSSIKRQIILCME ELDHTPDTSFERDVVCEDEDAFCLSLENIATL QKLLRQLEMQKSQNEAVCEGLRTQIRELW DRLQIPEEEREAVATIMSGSKAKVRKALQULE VDRLEELEKCKTMKKVIEAIRVELVQYWDQC FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR LKNYYEVHKELFEGVQKWEETWRLFLEFER KASDPNRFTNRGGNLLKEEKQRAKLQKMLP KLEEELKARIELWEQEHSKAFMVNGQKFME YVAEQWEMHRLEKERAKQERQLKNKKQTET EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK PVAASTCSGKKTPRTGRHGANKENLELNGSI LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS DSSTVGLQRELSKASKSDATSGILNSTNIQS
551	1901	A	4194	3	1008	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV VALLDICFILVAVPESLPEKMRPVSWGAQISW KQADPFASLKKVGKDSTVLL\ICITVCLSYLPE AG\QYSSFF\LYLR\QVIGFG\TVKLAAFIAMVGI LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML QLAWYGFGSQAWMMWAAGTVAAMSSTIFP AISALVSRNAESDQQGVAQGIITGIRGLCNGL GPALYGFIFYMFHVELTELGPKLNSNNVPLQ GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS SIWELSSFEEPGNQCTEL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide		1	USSN	location	corresponding	
	seq-	l				I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	I		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ľ	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1			peptide		/=possible nucleotide deletion, \=possible
				sequence	<u></u>	nucleotide insertion
552	1902	Α	4197	2	14302	ARPPPAPGSRQQKQKAAPGAAAAAELRGAR
						EPAPARRRGTMADGGEGEDEIQFLRTDDEVV
1]			LQCTATIHKEQQKLCLAAEGFGNRLCFLESTS
						NSKNVPPDLSICTFVLEQSLSVRALQEMLANT
1	i					VEKSEGQVDVEKWKFMMKTAQGGGHRTLL
						YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD
Í	1					VGLQEDTTGEACWWTIHPASKQRSEGEKVR
1						VGDDLILVSVSSERYLHLSYGNGSLHVDAAF
1						QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH
						GHMDECLTVPSGEHGEEQRRTVHYEGGAVS
						VHARSLWRLETLRVAWSGSHIRWGOPFRLR
i						[_ . . .
1						HVTTGKYLSLMEDKNLLLMDKEKADVKSTA FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS
1						
1						VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR
						KAIMHHEGHMDDGISLSRSQHEESRTARVIRS
						TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL
						SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR
1						QNLFQEEGMINLVLECIDRLHVYSSAAHFAD
1						VAGREAGESWKSILNSLYELLAALIRGNRKN
						CAQFSGSLDWLISRLERLEASSGILEVLHCVL
1						VESPEALNIIKEGHIKSIISLLDKHGRNHKVLD
}						VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL
						LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY
						YELMVDHTEPFVTAEATHLRVGWASTEGYSP
]						YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG
1						CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF
1						RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV
						RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL
1				ĺ	•	KVEHSREYKQERTYTRDLLGPTVSLTQAAFT
						PIPVDTSQIVLPPHLERIREKLAENIHELWVMN
	[- 1			[KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ
		- 1				ERNYNLOMSLETLKTLLALGCHVGISDEHAE
1		1				DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT
1		i	ı		ł	PSQEAMVDKLAENAHNVWARDRIRQGWTY
						GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS
			1	1	j	LREAVRTLLGYGYNLEAPDODHAARAEVCS
] .		I	ŀ		ł	GTGERFRIFRAEKTYAVKAGRWYFEFETVTA
			ļ			GDMRVGWSRPGCQPDQELGSDERAFAFDGF
		İ				KAQRWHQGNEHYGRSWQAGDVVGCMVDM
		1	1			NEHTMMFTLNGEILLDDSGSELAFKDFDVGD
]				GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC
	ł	ŀ	ł	ł		GLQEGYEPFAVNTNRDITMWLSKRLPOFLQV
		Ī		ļ	}	PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN
		ŀ			i	SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG
	ł		İ	Ì	1	LFGPKNDLEDYDADSDFEVLMKTAHGHLVP
		}			1	DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ
		1	l		l	RFLLRRTKPDYSTSHSARLTEDVLADDRDDY
]	ļ	.]	J	J	}	DFLMQTSTYYYSVRIFPGQEPANVWVGWITS
		1		İ	1	DFHOYDTGFDLDRVRTVTVTLGDEKGKVHE
		l			1	
; l			ļ			SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC
[[ĺ	ĺ	1	1	1	VVDAASGLLTFIANGKELSTYYQVEPSTKLFP
]]	J	J]	. 1	J	AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS
	ļ	ł	.	l	1	EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL
	i	İ			ļ	KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN
		ļ			į	RSVDILELTEQEELLKFHYHTLRLYSAVCALG
[. 1	į.	. [ſ	į	NHRVAHALCSHVDEPQLLYAIENKYMPGLLR
	l	.		1	ĺ	AGYYDLLIDIHLSSYATARLMMNEYIVPMT
	l			1	İ	EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF
1 .1	!	ł	i	ı	İ	SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE
L						AVKEGSLHARDPVGGTTEFLFVPLIKLFYTLLI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghtamic Acid,
nuci-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	•		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l i			peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1						MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
1						LEKELSVDDAKLQGAGEEEAKGGKRPKEGLL
1						QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI
						VAFSDDFVAKLQDNQRFRYNEVMQALNMSA
1						ALTARKTKEFRSPPQEQINMLLNFKDDKSECP
						CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG
1						NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES
1						DSKKSSTLQQLISETMVRWAQESVIEDPELVR
[]						AMFVLLHRQYDGIGGLVRALPKTYTINGVSV EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
						LGDIMNNKVFYQHPNLMRALGMHETVMEV
						MVNVLGGGESKEITFPKMVANCCRFLCYFCR
						ISRQNQKAMFDHLSYLLENSSVGLASPAMRG
) '						STPLDVAAASVMDNNELALALREPDLEKVVR
1						YLAGCGLQSCQMLVSKGYPDIGWNPVEGER
1			•			YLDFLRFAVFCNGESVEENANVVVRLLIRRPE
						CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
1		1				GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY
						SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
)						LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS
						AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL
						EVGFLPDLRAAASLDTAALSATDMALALNRY
						LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT
					•	VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS
1 1				}		MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH
1 1						YERCWKYYCLPGGWGNFGAASEEELHLSRK
1 1						LFWGIFDALSQKKYEQELFKLALPCLSAVAG
						ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK
1				ļ		LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE
1 1					ł	KEIYRWPIKESLKTMLARTMRTERTREGDSM
						ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
						NVTLSRDLHAMAEMMAENYHNIWAKKKKM
1			`			ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
						QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA
1 1		Í	ľ			YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF
						PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA
						SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
						RISLFGNDATSIVNCLHILGQTLDARTVMKTG
			1			LESVKSALRAFLDNAAEDLEKTMENLKQGQF
			1			THTRNQPKGVTQIINYTTVALLPMLSSLFEHI
1		ł	I			GQHQFGEDLILEDVQVSCYRILTSLYALGTSK
						SIYVERQRSALGECLAAFAGAFPVAFLETHLD
		- 1	ĺ			KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP
				ļ		SLEKLMEEIVELAESGIRYTQMPHVMEVILPM
		l				LCSYMSRWWEHGPENNPERAEMCCTALNSE
1 1	- 1	1	1	1	1	HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF
		- 1	1	!	1	SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
		l		ļ	ļ	SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
	ŀ	1				MVAEVFIYWSKSHNFKREEONFVVONEINN
		j		. [MSFLITDTKSKMSKAAVSDQERKKMKRKGD
-		ĺ	1	1	ľ	RYSMOTSLIVAALKRLLPIGLNICAPGDOELIA
		1				LAKNRFSLKDTEDEVRDIRSNIHLOGKLEDP
		1	ŀ			AIRWOMALYKDLPNRTDDTSDPEKTVERVL
			,		1	DIANVLFHLEQKSKRVGRRHYCLVEHPORSK
}	J	j		•]	J	KAVWHKLLSKORKRAVVACFRMAPLYNLPR
	.			ļ	İ	HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA
			•	ļ		KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT
				<u> </u>	·	EKCKLEEDFLYMAYADIMAKSCHDEEDDDG
				· · · · · · · · · · · · · · · · · · ·		·

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						EEEVKSFEEKEMEKQKLLYQQARLHDRGAA EMVLQTISASKGETGPMVAATLKLGIAILNGG NSTVQQKMLDYLKEKKDVGFFQSLAGLMQS CSVLDLNAFERQNKAEGLGMVTEEGSGEKV
				i i		LQDDEFTCDLFRFLQLLCEGHNSDFQNYLRT QTGNNTTVNIIISTVDYLLRVQESISDFYWYY SGKDVIDEQGQRNFSKAIQVAKQVFNILTEYI
						QGPCTGNQQSLAHSRLWDAVVGFLHVFAHM QMKLSQDSSQIELLKELMDLQKDMVVMLLS MLEGNVVNGTIGKQMVDMLVESSNNVEMIL KFFDMFLKLKDLTSSDTFKEYDPDGKGVISK
						RDFHKAMESHKHYTQSETEFLLSCAETDENE TLDYEEFVKRFHEPAKDIGFNVAVLLTNLSEH MPNDTRLQTFLELAESVLNYFQPFLGRIEIMG SAKRIERVYFEISESSRTQWEKPQVKESKRQFI
			,			FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA QISESDLNERSANKEESEKERPEEQGPRMAFF SILTVRSALFALRYNILTLMRMLSLKSLKKQM KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV
			,			ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL ANMPDPTQDEVRGDGEEGERKPLEAALPSED LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK
						EEEKEEKEETKSEPEKAEGEDGEKEEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYNMRMLALFVAFAINFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS
	•					SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF
				•	,	VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT
553	1903	A	4199	31	767	LPELNGRGAGLRRAEPSERGGGAERTQQVAA LPLSHGHSHGGGGCRCAAER/VGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW
			·			ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK GIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI
554	1904	A	4200	1	961	HISKNFGADITKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL
	130.		.200	•		EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT
						WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDOOKLELKSPVLRKOACHOWKHSEVESGV
						LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGT\ RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH
555	1905	A	4211	331	2419	KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE
						QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNP NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPLS DSNRDHTANRQQRST\SPVARRTRSQTSVNFN
<u> </u>						GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNGI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end mucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GGAAGIPRANASRTNFSSHTNQSGGSELRQRE GQRFGAAHVWENGARSNVTVRNTNQRLEPI RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV QQTTRRSVRRRGRTRVFLEQDRERERRGTAY TPFSNSRLVSRTTVEEGEESSRSSTAVRRHPTTT LDLQVR\RIRPGENRDRDSIANRTRSRVGLAE NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP LRRISENELVEPSSVALRSILRQIMTGFGELSSL MEADSESELQRNGQHLPDMHSELSNLGTDN NRSQHREGSSQDRQAQGDSTEMHGENETTQP HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH
						FFLLNESDDDDRIRGLTKEQIDNLSTRHYEHN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HIHCIDRWLSENCTCPICRQPVLGSNIANNG
556	1906	Α	4212	3	462	LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR KSPENTEGKDGSKVTKQEPTRRSARLSAKPA PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK QEAGKEGTAPSENGETKAEEIHISRSTVNVST SRGTPPSTLSVKGQIETVRVKGTEN
557	1907	A	4213	774	507	ARRESCLTLQTSWGHRHGPPRPANFVFLVET GFLHIGQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLIDSCKNFIVYKIMYTL
558	1908	A	4225	3	1253	TYRHAEREHPETSSATKVSYDYRHKRPKLLD GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKSSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGITLHERFSVKMADIHKADV NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP
559	1909	A	4235	ì	323	KFSIPFFLRWSFTLV\PRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS
560	1910	A	4246	2	1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLWDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHTYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD
561	1911	A	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	i	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	-			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	Ì			peptide		/=possible nucleotide deletion, \=possible
				sequence	L	nucleotide insertion
						INTLEPVEEDAEMRGDQGTPTTGLLMIVLGLI
						FMKGNTIKETEAWDFLLAL\GVYPTKKHLIFG
						DPKKLITEDFVRQRYLEYRRIPHTDPVDYEFQ
		i				WGPRTNLETSKMKVLKFVAKVHNQDPKDW
						PAQYCEALADEENRARPQPSGPAPSS
562	1912	Α	4260	1	1498	MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF
						WLHARLQKCFLSRGCGSYCAGAKASPLPGK
		ļ				MAMGLMCGRRELLRLLQSGRRVHSVAGPSQ
i l	ļ					WLGKPLTTRLLFPAAPCCCRPHYLFLAASGPR
		İ		!		SLSTSAISFAEVQVQAPPVVAATPSPTAVPEV
		1				ASGETADVVQTAAEQSFAELGLGSYTPVGLI
	1	l	1			QNLLEFMHVDLGLPWWGAIAACTVFARCLIF
('	1	1				PLIVTGQREAARIHNHLPEIQKFSSRIREAKLA
	·	1				GDHIEYYKASSEMALYQKKHGIKLYKPLILPV
	•	1				TQAPIFISFFIALREMANLPVPSLQTGGLWWF
<u> </u>		1				QDLTVSDPIYILPLAVTATMWAVLELGAETG
						VQSSDLQWMRNVIRMMPLITLPITMHFPTAV
i ·						FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR
						VVHDLDKLPPREGFLESFKKGWKNAEMTRQ
						LREREQRMRNQLELAARGPLRQTFTHNPLLQ
	1010		10.65	(00	117	PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG
563	1913	Α	4265	623	116	MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQN
i l		İ				THGTRRTLSSGPSNNLPLPLSSSATMPSMQCK
						HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGV\L
		ł			·	PRGSGNPPHGTSILTAPPALLPHPPTHPTQQSF
i I	•					LIQENNNTNHTHSHTHTYTETLSFFLYICVNN
564	1014		4070		2/0	DRMEWGKSVF
564	1914	A	4270	3	368	ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL
1						GCTWCLGLLQVGPAAQVMAYLFTIINSLQGF
(ĺ	[FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES
	1016		4000	02	406	ETYTLSSKMGPDSKPSEGDVFPRTSE
565	1915	Α	4288	83	406	RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS
i		ĺ		İ		PPSALLAPTKPRALGTLRLYECSPELCTTMLP
			•			PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP
566	1916	Α .	4298	1041	229	GOTGASRTPRT
200	סובו	^ .	4278	1041	449	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL
						GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS
						GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR
						VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA
						DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR
						VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
						FHLGASEWTPRLLIWR
567	1917	A	4299	1	1106	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
307	171/	Λ.	4677	1	1100	
[DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WETVEKNEPAIGGTGPTSDTGWGCMI PCGO
						WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
						VLNAFIDRKDSYYSIHOIAOMGVGEGKSIGO
						WYGPNTVAOVLKKLAVFDTWSSLAVHIAMD
					İ	•
				1		NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD
				l		
				l		INEAYVETLKHCFMMPQSLGVIGGKPNSAHY
		l		-		FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF
1			, ,			
560	1010		4300	2012	1942	GAECCLGMTRKTFGFLRFFFSMLG
568	1918	A	4300	2012	1843	GAECCLGMTRKTFGFLRFFFSMLG SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS
568 569	1918	A	4300	2012	1843	GAECCLGMTRKTFGFLRFFFSMLG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI CVWWGVGCVKCLPRAHHIWQEKPLGPHRT VTESKLEAEGKTKEKAREKERKKS
570	1920	A	4308	3		RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS GKRNKLRVYYLSWLRNKILHNDPEVEKKQG WTTVGDMEGCGHYRVVKYERIKFLVIALKSS VEVYAWAPKPYHKFMAFKSFADLPHRPLLV DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE KEHREFRAKTNRDLEIKDQEIEKLRIELDESK QHLEQEQQKAALAREECLRLTELLGESEHQL HLTRQEKDSIQQSFSKEAKAQALQAQQREQE LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT KLKEECCTLAKKLEQISQ
572	1922	A	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLYT DINEAYVETLKHCFHGWPQFPG/VVHREGK PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS LFLTIPNLAISWEGHIVVYSSTEEKTTLKERM HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSINGLHISLINPLAMRLPIHCVCVT KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD FSKGSK
574	1924		4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK LPSARAKIRITSSPIFITFYILVFVVALVGIARA VVSMTVSTSNAATVADKILWEITRFFLLAIEL SVIILGLAFGHILESKSSIKRVLAITTVLSLAYSV TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY VYAGILALLNLLQGLGSVLLCFDIIEGLCCVD ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	Ā	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM KIIKGSSGTPKLSYTGRDDRHFVPMGLYTVRT VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid	to last amino acid residue of peptide	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
		•		residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
577	1927	A	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW ASSNSRASTPKVAGITGARPHARIIFVFLIEMG FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	A	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW SGWSRTPDLR
579	1929	A	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP CWPGWSSTPDLK
580	1930	A	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ \VFKKGNIHILHELFQNKEEGAFPNS/FYEASFT LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK QLKSSDL
581	1931	Α	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS RLPHATIVDHRPQHRWLETCNAPPQLIQGKA RSAPKPSQASGHFSVELVRGYAGFGLTLGGG RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD LVLHINGESTQGLTVHAQAVERIRAGGPQLHL VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG SPE
582	1932	A	4424	194	449	VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE LEQELLEHGRDAASVQAATSVQAMQGKTTL PS\QGPLQRPSRLVFT\DVANAIHV
583	1933 .	Α	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV SAPPALLQDTSV
584	1934	Α	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT APATQHSQAGPATGQAYGPHTYTEPAKPKK GQQLWNRKPAPGTLEVSSSTSRSDPLLLPPR ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR GNLSGKPDDWP/LGHERVCGALLHRL*VGGG QGPHGKAAQGGAAGAAAGRLGLYH
585	1935	Α	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT SIFDDFAHYEKRQ
586	1936	A	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI TVQSIVIQSLNKTLTRREDIDVLQPTLVNAGH FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK LRLTGALPCQLVAQKVKSLLWGQGFPDYVA PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI SSSFPEANSGNERTILISTAVTFVDVSAPAEAG FRAPPAINARLPFNFFFPFV
587	1937	A	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP CPANFCIII/DFLVETGFHHVGQASHELLTSGD PPTSASQSAGITGMSYHTWFGES
589	1939	A	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC PPVELPWAPRRGHRLSPADDELYQRTRISLLQ REAAQAMYIDSYNSRGFMINGNRVLGPCALL PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
1		l ·		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ĺ	ĺ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide		/=possible nucleotide deletion, \=possible
İ	i	ŀ		sequence		nucleotide insertion
 		 	 			VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ
İ		1				DTPNACATFNFLCHEGRVTGAALIPPPGGTSL
		l			ì	TSLGQAAQ
590	1940	Α	4492	1	472	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
1 390	1,770	^	4472	1	7/2	PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP
1 .	}	ļ	l			RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR
1				i		ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS
1						
501	104		4405	1444	1116	VPHQGGLPGPIRVAPSSAGQREASQGPPGR
591	1941	Α	4495	1444	1116	LAARFTLAKTWNQLKRP\TMIDSIKKTR\YTYT
Į.						MEYYADTERNEIMSF\AGTWVELEAIILSKLM
	}		l			LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL
	<u> </u>					EPWDSSCFPHPSSGV
592	1942	Α	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
			[VCGGLSLLANAWGILSVGAKQKKWKPLEFL
		1	1	1		LCTLAATHMLNVAVPIATYSVVQLRRQRPDF
				Į.		EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR
				Í		MWMVCWPVNYRLSNAKKQAGHTVMGIWM -
				}		GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
1						GLGFGVCFLLLVGGSVAMGVICTAIALFQTL
				l		AVQVGRQADHRAFTVPTIVVEDAQGKRRSSI
1			1	į ·		DGSEPAKTSLOTTGLVTTIVFIYDCLMGFPVL
			ļ			GPFSLADTHLSDLPYTWGDRDSGGACVM
593	1943	A	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC
	· -					HFPASASQVAGTTHARHHTQLIF\AFLVENGL
		1				C
594	1944	Α	4507	1327	647	KMAGGVRPLRGLRALCRVLLFLSOFCILSGG
1		**	1007	132'	* ''	ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS
				ĺ	·	CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII
		ļ		_		NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ
				•		RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
į .						GGYKWVYGLWLLRHHPRWGLGADRF\YLGP
						VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG
İ		İ				PADGSLYI
595	1945		4512	533	264	FFFKMESYSVARLECSGAISAPCNLHLLGSNN
293	1945	Α	4512	333	204	
		İ	İ '	i		SPASASRV/AGNIGARHHTQQIFVLLVQMRVH
						YVGQDGLDLL/NLMIHPPRSPKVLGLQA
596	1946	Α	4513	3	1674	HASDHLYPNFLVNELILKQKQRFEEKRFKLD
						HSVSSTNGHRWQIFQDWLGTDQDNLDLANV
						NLMLELLVQKKKQLEAESHAAQLQILMEFLK
						VARRNKREQLEQIQKELSVLEEDIKRVEEMS
						GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP
					,	PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE
						DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS
			1			KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV
1						FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA
						VDIHYPENEMTCNSKISCISWSSYHKNLLASS
]						DYEGTVILWDGFTGQRSKVYQEHEKRCWSV
						DFNLMDPKLLASGSDDAKVKLWSTNLDNSV
						ASIEAKANVCCVKFSPSSRYHLAFGCADHCV
						HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG
						EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN
1						EKNFV\GLASNGDYIACGSENNSLYLYYKGLS
						KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV
[[!			CWRALPDGESNVLIAANS\QGTI\KVLELV
597	1947	A	4518	536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS
""	174/	Α.	4710	J30	024	ASQVAEITSVRHHTWLIFCILLGOMGFHHVGE
[$I \cap A \cap I \cap I \cap C \cap I \cap C \cap A \cap I \cap C \cap A \cap C \cap A \cap C \cap A \cap C \cap A \cap C \cap A \cap C \cap A \cap C \cap C$
500	1040	•	4524		204	QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP
598	1948	A	4524	1	384	QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ : ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RRREMQSQSVMLALRRGDAVWLLSHDHDG
				•		YGAYSNHGKYTTFSGFLVYPDLAPAAPPGLG ASELL
599	1949	A	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI PPPWLPKVLGLQA
600	1950	A	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS DPPASASRVAGTTGARHHTQLIFVFLVETGFH \MLARDGLKLLTSSDFPASASQSSWDYRREPP RLANFFVFLVETGSRYVAQAGVQWLFTGAIP LLISTGVLTCSVSDLGRFTPP
601	1951	A	4533	1460	403	HEVQESIHIFLESEFSRGISDNYTLALITYALSS VGSPKAKEALNMLTWRAEQEGGMQFWVSSE SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE GIPIMRWLSRQRNSLGGFASTQDTTVALKALS EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT HNRLLLQTAELADGTANGSV/SISANGFGFAI CQLNVVYNVKASGSSRRRRSIQNQEAFDLDV AVKENKDDLNHVDLNVCTSFSGPGRSGMAL MEVNLLSGFMVPSEAISLSETVKKVEYDHGK LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD VQRLPSL
602	1952	A	4540	1963	295	MRAPGRPALRPLPLPPLLLLLLSSPWGRAVPC VSGGLPKPANITFLSINMKNVLQWTPPEGLQG VKVTYTVQYFIYGQKKWLNKSECRNINRTYC DLSAETSDYEHQYYAKVKAIWGTKCSKWAE SGRFYPFLETQIGPPEVALITDEKSISVVLTAP EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT KSNRTWSQCVTNHTLVLTWLEPNTLYCVHV ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HPANLILIYGNEFDKRFFVPAIEKIVINFINTL NISIDDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENTN EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEBQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN
603	1953	A	4543	3	600	YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV
604	1954	A	4548	3	938	QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA VKTVGSVVSSVALTGVLSGNGGTNVNMPVS KPTSWAAIASKPAKPQPKMKTKSGPVMGGG LPPPPIKHNMDIGTWDNKGPVPKAPVPQQAP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ
605	1055		4552	1	2204	QPPQ
605	1955		4553	2	2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE GPGLGALDRLRAHASAMGDEDLPGMAALQP HGVPGDGEGPHERGPPPASAPVGGTVTLRED SAKRLERRARRISACLSDYSLASDSGVFEPLT KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN LADYDSLSEMQLRWHSVQVFTSLNHQGRGR LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR TTAQLQAVERELAEERAKLEYTEEEVLEMER KEEQAEAISERSWQADSVDSGCSNCTQTSPPY PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL KVDKETNTEDLFLEEAASLVKERPSRRARGSP FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS STLPRKSPFVRNTLERTTLRYKQSCRSSLAEL MARTSLDLELDLQASRTRQRQLNEELCALRE LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR EAERQTRQTKLDYRHEQAAEKMLKKASKEI YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL PADDV
606	1956	A	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP VLLLQDSSGDYSLAHVREMACSIVDQKFPEC GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTPVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	A	4563		4499	SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI QPGAFRLRNLNTLLLNNNQIKRIPSGAFEDL ENLKYLYLYKNEIQSIDRQAFKGLASLEQLYL

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Laucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first amino acid	acid residue of peptide	Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
l		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	sequence	/=possible nucleotide deletion. \=possible
]	J .		sequence		nucleotide insertion
					<u> </u>	HFNQIETLDPDSFQHLPKLERLFLHNNRITHL
ļ		ľ	1			VPGTFNHLESMKRLRLDSNTLHCDCEILWLA
1	1	1	1]	DLLKTYAESGNAQAAAICEYPRRIQGRSVATI
İ						TPEELNCERPRITSEPQDADVTSGNTVYFTCR
	ļ.				ļ	AEGNPKPEIIWLRNNNELSMKTDSRLNLLDD
	1	l			}	GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ
						EVTLRYFGSPARPTFVIQPQNTEVLVGESVTL
Ī	1	ļ				ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS GGLYIQNVVQGDSGEYACSATNNIDSVHATA
	}	Ì				FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK
1						GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI
		1	1			SGVALHDQGQYECQAVNIIGSQKVVAHLTVO
			[PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP
J	}]]			EPAITWNKDGVQVTESGKFHISPEGFLTINDV
						GPADAGRYECVARNTIGSASVSMVLSVNVPD
Ī	1	}				VSRNGDPFVATSIVEALATVDRAINSTRTHLF
		j				DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF
1						ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS
Ì			[·			POYLNLIANLSGCTAHRRVNNCSDMCFHQKY
						RTHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI
ļ						GTETVTPDEQFTHMLMQWGQFLDHDLDSTV
			1			VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN
						DSRARSGARCMFFVRSSPVCGSGMTSLLMNS
Ì	ŀ					VYPREQINQLTSYIDASNVYGSTEHEARSIRD
						LASHRGLLRQGIVQRSGKPLLPFATGPPTECM
			,			RDENESPIPCFLAGDHRANEQLGLTSMHTLW
]					FREHNRIATELLKLNPHWDGDTTYYETRKIVG
						AEIQHITYQHWLPKILGEVGMRTLGEYHGYD
	1		i			PGINAGIFNAFAT\AAFRFGHTLVNPLLLPGLD ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL
						RGLFGVAGKMRVPSQLLNTELTERLFSMAHT
						VALDLAAINIQRGRDHGIPPYHDYRVYCNLS
	1 1					AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID
•						LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR
						DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
	[' i	1	-		NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD
	1					LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE
		ļ				FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS
		- 1	1	ſ	ĺ	A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTKWK
		ł	}	1	ł	KDACTICECKDGQVTCFVEACPPATCAVPVNI
			ŀ	ļ		PGACCPVCLQKRAEEKP
608	1958	A	4566	354	1135	FSFLC/GVSGRLGLDSEEDYYTPOKVDVPKAL
		ľ	ł	1	i	IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL
		1	1	•		GLNQCMSGIINHEAYHEVPYTTSFTLAKOLSF
	.	I		l		YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGO
		1		ľ		LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD
				1		EFTIAATDDNHIFAWGNGGNGRLAMTPTERP
		-		1		HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI
	. [1		Ī	[VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGE GGPDAW
609	1959	A	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
.				-	• • • • • • • • • • • • • • • • • • • •	PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF
		l		1		HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH
		- 1	1	l	!	RARPRINLRNVIYSFAVTYCLNYISLAMSSTL
		[[Ī	KLSFHVLSGS
610	1960	A	4570	697	467	ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT
i					1	GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N
		- 1				LVIRPPRPPKVLGLQA

SEQ II	D SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of		hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
[amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion_\=possible
Ì	1	l		sequence	}	nucleotide insertion
611	1961	A	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
					l	WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
	j		J	J		LCAATAVLLSAQGGPVQSKSPRFASWDEMN
1		f				VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
						LSACGSACOGTEGSTDLPLAPESRVDPEVLHS
	1	ì		1		LQTQLKAQNSRIQQLFHKVAQQQRHLEKQHL
1		ľ	[1	RIQHLOSOFGLLDHKHLDHEVAKPARRKRLP
						EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
	1 1					QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIOR
]			RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
						EKVHSITGDRNSRLAVQLRDWDGNAELLOFS
1	1 1				{	VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
1						LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
1						GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
1					ļ	WRGRYYPLQATTMLIQPMAAEAAS
612	1962	A	4575	162	3	FFFETESRSVAQAGVQWRDLSSLOPPPPG\SR
]	1,500	••		102	١	GSPASASPVAGITGTRHHRTRG
613	1963	Α	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
	1,500	· •	.50,	00,	321	SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
l			1		1	HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
				•		LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	Α	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
] 014	1704	^	4367	121	233	GGLFCAWVGTILLVVAMATDHWMQYRLSGS
				,		FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
						FMILSALCAISGIIMGIMAF/GWVAVLMTFFA GIFYMCAYRVHECRRLSTPR
615	1965	A	4590	2	414	
1 013	1903	Δ.	14390	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
				-		ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHSGHQEQLNRKR
)					ERRPLPKNARPSPWVPALADEWNTLHQEVTT
						TRLPAGSOEPVKD
616	1966	A	4592	773	488	
010	1,000	^	4372	113	400	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
	- 1 - 1					
617	1967	В	4595	84	478	AGLELPTSGDPPALASQSAGITGVTTVPSGPG
017	1907	Б	4393	04	4/0	XRHGLREPLLERRCAAASSFQHSSSLGRELPY
	- 1 - 1		i l			DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT
İ			1			REKQLQELQQQEEEERQRQQRREERRQQNL
	j l					RARSREHPVVGHPDPALPPSGVNCSGCGAEL
618	1050	A	1504	2045	1100	HCQDAR*
010	1968	A	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
	1					SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
ļ						MVGNETTYEDGHGSRKNITDLVEGAKKANG
	l l					VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA
						MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF
						HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV
į						YLHLRQTWLAFMULSILEVIIILLLIFLRKRILI
			 	ļ		AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
						AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
						CNPETFPSSNESRQCPNARCQFAFYGGESGYH
		- 1			·	RALLGLQIFNAFMFFWLANFVLALGQVTLAG
1] [- 1		1	İ	AFASYYWALRKPDDLPAFPLFSAFGRALRYH
		- 1				TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN
		l				KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
				ļ	·	IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
]]	ļ				TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
	1 {	1	. 1			APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
	į t	,				
						VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL
619	1969	A	4601	2	357	VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i		İ		peptide		/=possible nucleotide deletion, \=possible
1	İ	i	1	sequence	ľ	nucleotide insertion
						GLAYGHFSYEFSNHRDVVVDLQGWVTGNGK
ł	i	l	1			GLIYLTDPQIHSVDQKVFTTNFGKRGIFYFFN
1			ŀ			NQHVECNEICHRLSLTRPSMEKPCKS
620	1970	A	4606	1	2415	MERLWGLFQRAQQLSPRSSQTVYQRVEGPR
		1	.000	•	2413	KGHLEEEEDGEEGAETLAHFCPMELRGPEP
	Ì	İ				LGSRPRQPNLIPWAAAGRRAAPYLVLTALLIF
1	ĺ	[.			1	TGAELI GVIAPROCOO ACCRONIATIONNA
1]	TGAFLLGYVAFRGSCQACGDSVLVVSEDVN
1	ĺ	1			ľ	YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
		ł				EDTIRQTSLRERVAGSAGMAALTQDIRAALS
1		ł	1 .		1	RQKLDHVWTDTHYVGLQFPDPAHPNTLHWV
1						DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL
J .		ļ]	VYAHYGRPEDLQDLRARGVDPVGRLLLVRV
		į	1		•	GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ
						DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF
1 1					Í	NQTQFPPVASSGLPSIPAQPISADIASRLLRKL
1						KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN
1 1					Ì	NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA
1					ļ	WGPGAAKSAVGTAILLELVRTFSSMVSNGFR
1 1					i	PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL
1 1						HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL
1 1			j l			IESVLKQVDSPNHSGQTLYEQVVFTN\PSWD\
						AEVIRPLPM\DSSAY\SFTAFVGVPAVEFSFME\
1 1						DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA
[[QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV
			į			LRHIGNLNEFSGDLKARGLTLQWVYSARGDY
1 1	-		l i			IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV
	1					EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL
1 1	ł					DHLRLLRSNSSGTPGATSSTGFQ\ESRFRRQL\
			·			ALL\TWDACKGAANALSGDVWNIDNNF
621	1971	A	4610	793	334	ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\
				1		NTLVLKQQTFIESARSIGASDMTVLLRHILPGT
			}	1		GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP
1 (1			ĺ		EWGAMLNEARADMVIAPHVAVFPALAIFLTV
						LAFNLLGDGLRDALDPKIKG
622	1972	A	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ
1	ł		ŀ		l	CTIACRGKNMEVRLIFLSGLCLAVAVVWAVF
1 1	1	- 1				RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK
1						SCVILLGLLLYDVFFVFITPFITKNGESIMVEL
i J		ļ	1			AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV
1	ł		ŀ			TITOT TOTAL TERROR D. COLLOS A COLLOS DE COLLO
]	1	ļ				VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL
1	İ		- 1	1		TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV
		1	į			AWETVREMKKFWERVTS
623	1973	A	4619	17	691	TLVSVVEFVRRADLTREDLAPSSVDSGQAGF
				•	- ·	
			- 1	ļ		GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT
	1		i			DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ
	ľ	1	i	Í	Ĭ	IGHFLCLVILVYCAEYINEAAAMNWRLFSKY
	4		- !	İ		QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW
	1	ł	ĺ	ł	l	KTLNVMTDLKNAQERRKEKKRRKED*GAA
	ŀ		1			AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG
624	1074	<u> </u>	4600			LKNRCFI
624	1974	A	4622	164	668	VSCYTALQSIMNQPESANDPEPLCAVCGQAH
		- }	1		ŀ	SLEENHFYSYPEEVDDDLICHICLQALLDPLD
	- 1		- 1		1	TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL
		- 1			ļ	QHCKKSSILVNKLLNKLLVTCPFREHCTQVL
	1	l		i	ł	QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED
						CLSPGVHHCSEV
625	1975	A	4625	474	473	CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP
						PPLLIPSS*LSP
					<u>-</u> -	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methiomine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
626	1976	A	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP ASASQVAGIAGTHH
627	1977	Α	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP ARAYL
628	1978	A	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK VFYICFTEFLLFLYFL*LFIIKVSCSII*CSTICVF SYKSFAVIIFFVDNTRFFSFGF
629	1979	A	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH PKLVFSQEGRYVKNTASASSWPVFSSAWNYF AGWRNPQKTAFVERFQHLSCVLGKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KMSPDVGIWAIYWSAAGYWPLIGFPGTPTQQ EPALHRVGVYLDRGTGNVSFYSAVDGVHLH TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL
630	1980	A	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG FHDVGQDGLDLLTS*STPSASQSAEITGVSHC TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ NPVFLERRPRALHSSPGLTTQRILWAQGLWV GAGSTGCSRGPRGEGVFREG
632	1982	A	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	A	4696	1	1365	RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFTTYRITKTSRG EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP
634	1984	Ā	4708	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM
635	1985	Α	4709	42	341	FPHAIGGSCRAA*RRRKTLQFPCYH YIKQPDAKERRRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED
636	1986	A	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS RSTTVTVA*LMQKLNLSMNDAYYIVIMKMSS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	sco-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1			peptide	l •	/-possible nucleotide deletion, \-possible
1	•		ŀ	sequence		nucleotide insertion
			<u> </u>			ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK
		1			ļ	MYFTTPSNHNAYQVDSVQST
637	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH
1						LKSESFOVLTLSPRLEFSGLISAHCNLRLPGSS
		1				DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
		1				HAG*AGLELLTSGDPPASASRSAGITGVSHHA
1		1				RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
1	, ,,,,,	1	'''		***	TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
			·			YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL
						GQYTSQGGVTAWRKICPIFEGIGYASQMIVIL
1						LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
j		}			}	WNTEHCMEFOKTNGSLNGTSENATSPVIEFW
639	1989	A	4743	1040	699	QGLTLLPRMECSATITAHCSLELPGSIDLPTSA
""	'''	1	'' '	2010	""	S*VARTTGTHHHPWLILVLLL*TWGSYYVAO
						AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA
		1				TSNHVLYTQEGLRRGKEG
640	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
1 010	1550	 ^ _	*//1	327	*	WIRRPCLPSGCLKMNREIGPLOHSLCCPGWS
1						QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA
1		ļ	1		}	MARSRLTATSASQVQAILLPQPPGTTDSCSPS
ŀ		,				PDHEQQPLSWVLPPPQKDMNPREQQVALGP
		1				QAAALPWAVWRNDCFPR
641	1991	A	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
1	1,7,7,1	1 **	1,,00		7/3	LQLAASPYFSPSWAECPQPVPAGTHATWCLA
1		1				RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
i i	·	i '				FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
		ŀ				QRDTDTGVHTGSGTHTHAHTPPEK
642	1992	A	4798	1	487	GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
) 0.2	1772	١.,	1 1//0	•	707	FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
				·		TWWFGVKFAAGGLGTFHALLNTAVHVVMY
				-		SYYGLSALGPAYQKYLWWKKYLTSLOLVQF
		ļ				VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
						FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
1			''''	_	371	QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
}						MVYFVGENNGDSSHNPVLAATGVGLDVAQS
,			! !			WEKAIRQALMPVTPQASVCTSPGQGKDHSK
						Q*ASVCTSPGQGKDHSKQ
644	1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
			''''			LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
						AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
						SYKDIWGWPCLCGVLHAYIPLLV
645	1995	A	4805	458	126	LLWTTVLCOTPARPOSTMIHLGHILFLLLLPV
• • •			.505			AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
1 1		' :	¦	i	i	PLLAGLVAADAVASLLIVGAVFLCARPRRSP
				ļ	 	AQEDGKVYINMPGRG
646	1996	Α	4817	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
""	1,,,,	41	101/	"	כנטו	LOPOAPGHDMTSIPFPGDRLLOVDGVILCGLT
					ļ !	HKQAVQCLKGPGQVARLVLERRVPRSTOOC
1				ĺ		PSANDSMGDERTAVSLVTALPGRPSSCVSVT
]					İ	DGPKF*SSN*KRIANGLGPSFVQMEKESCSHL
					i	KSDLVRIKRLFPGHPAEENGALAAGDIILGRE
	İ			ļ	ļ	WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
				j		YYPAAVEVLHILLRGAPQEVILLLCRPPPGAL
}				-		PELEQEWQTPELSADKEFTRATCTDSCTSPIL
						GSRQLGGTVPPQMQGKAWGLRPESSQKAIR
			 			EGTMGAKTERDLGPVP
647	1997	A	4854	1044	335	PRVRGDWPLEKKKSNSNIHPIFSWCGSTDSKD
<u> </u>	2771	л	T0.74	1044	ردد	TIVANOD MATERIAN PROBABILISMO PROPERTY AND PROPERTY OF THE PRO

No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Docation Docation	1	NO: of	hod	ID NO:			D=Aspartic Acid, E=Glutamic Acid,
1998 A 2030 837 A 2030 A							F=Phenylalanine, G=Glycine, H=Histidine,
1914 ng to first amino aeid residue of peptide sequence 1915 191							
mimo acid residue of peptide sequence Ti-Threosine, V-valine, W-Typiophan, Y-Typiophan, Y-Sup codon, Peppide sequence Sequence Ti-Threosine, V-Valine, W-Typiophan, Y-Typiophan, Y-Typionia, X-Valine, W-Typiophan, Y-Typionia, X-Valine, W-Typiophan, Y-Typiophan, Y-Typionia, X-Valine, W-Typiophan, Y-Typiophan, Y-Typionia, X-Valine, W-Typiophan, Y-Typionia, X-Valine, W-Typiophan, Y-Typioph		ucilic					
Pepidid	1 4414	ľ	Ì	7.7			
Peptide	1						
					peptide	,	
GPPWESKNSTA-WRGRDSREREL ELVÄLSRK HPELIDA-TINFFIKHEDEL YQPIVKHISFD FFKHKYQNIDGTVAAYRLPYLLVODSVVLK QDSIYYEHYYELQPWKHTPYKSI,SDLLEK LKWAXDHDEEAKIAKAQQEFARNNIMGD DIFCYYFQTFPRMPYKY ACMIEAVCSADEEEDFAEEDCPELVPAETTO ACMIEAVCSADEEEDFAEEDCPELVPAETTO SEEEKSGLGAKIPYTIITOYLGAGKTTILNYI LTEQHSKRVAYINDFQEGSALPKSLAVSQG GBLYEEVLELRINGCLCCSVKDMGLRAIENLM QKKGFYPLLETTICAPFGAVASMWVDA ELGSDIYLDGIITIVDSKYGLKH.AEEKPDGLI NEATRQVALDALLNYGTUQVIGHELYDKKLRY TIRSINGLQQLETQRSRVDLSNYLDLHAFDSL NEATRQVALDALLNYGTUQVIGHELYDLETPV SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKQLFLAY TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKAP TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKAP TYTETEKQWTHEREDQVCT SWKDDITETRIR.VLLGRNLDKDLKAP TYTETEKQWTHEREDATA TYTETEKQWTHEREDQVCT TYTETEKQWTHEREDATA TYTETEKQWTHEREDA					sequence	ļ	nucleotide insertion
HPELIDAATINFFFKHIDENLYGPIVKHISFFD							IVMPTYDLTDSVLETMGRVSLDMMSVQANT
	i		}				
				1			
LKWAKDHDEAKKIAKAGQEFARNILMGD	ŀ						
	ŀ			1			1
1998			ĺ	[•) _
SEEEEKS.G.G.AKIPYTIITGYI.G.AGK.TIL.IMY	648	1998	A	4867	2030	837	
LTEQHSKRYAVILNEGEGSALEKSILAVSOG GELYEEWLEINGCLCCSVORGRAIENLM QKKCKFDYILLETGLADPGAVASMFWVDA ELGSDIYLDGITIYDSKYGLIKLEEKPDGLI NEATRQVALADALLNKTDLVPEEDVKKLRT TIRSINGLQGLETQRSRYDLSNVLDLHAFDSL SIGISLQKKLQHVRGTQPHLDGYTITTDVPG NAKEHINNFQNLLWEKNVRNKDNHCMEV RILKGLVSKKDSQQVIVQQVHELYDLEETPY SWKDDTERTNRLVLLGRILDKDLKQLFLAT VTETEKQWTIMFKEDQVCT				'''	2050	105.	
GELYEEWLELRINGCLCCSVKDNGLRINENM QKKGKPYDILETTGLAPPOAVASMFWVDA ELGSDIVLDGIITIVDSKYGLKHLAEEKPDGLI NEATRQVALADALLNKTDLVPEEDVKKLRT TRISINGLQGILETQRSRYDLSNVLDLHAFDSL SGISLQKKLQHVRGTQPHLDQSIVTITDVPG NAKEEHLNMFQNLLWERNKNRKDNHCMEW IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV SWKDDTERTINRLVLLGRINLDLKQLFLAT VTETEKQWTTHFKEDQVCT VTETEKQWTTHFKEDQVCT QAGLELRTSGDPFASASQSAGITGVSHLA*P TYPETCH STORM CONTINUE TO THE CONTINU			İ	}			
ELGSDIYLDGIITIVDSKYGLKH.AEEKPOCLI NEATRQVALADALINKIPPEDVSKLRT TIRSINGLGQILETQRSRVDLSNYLDLHAFDSL SGISLQKKLQHYPGTQPHLDQSIYTITDYPG NAKEEHLNMFQNLL WEKNYRNNCDNHCMEV RI.KGLVSIKDKSQQVIVQGVHELYDLEETPY SWKDDTERTRINLVLIGRIDLKDDILKQLFIAT VTETEKQWTTHFKEDQVCT	1		1				•
NEATRQVALADALINKTDLVPEEDVKKLRT			ĺ	Í			QKKGKFDYILLETTGLADPGAVASMFWVDA
TIRSINGLQQILETQRSRVDLSNYLDLHAFDSL							
SGISLOKKLÖHVPĞTQPHLDQSIVTITPDVPG NAKEEHLNMFIQNILLWEKNVRNKDNHCMEV IRLKGLVSIKDKSQQVIYQGVIELVDLEETPV SWKDDTERTNRLVLLGRNLDKDLKQLFIAT VTETEKQWTHFKEDQVCT			1				
NAKEEHLNMF1QNLLWEKNVRNDNHCMEV RILKGL VSIKDK SQQVTQVEHE TPU EETPV SWKDDTERTNI.VLLGRNLDKDILKQLFIAT VTETEKQWTTHFKEDQVCT				1			
					•		
SWKDDTERTNRL VLLGRNLDKDILKQLFIAT							
VIETEKQWTTHFKEDQVCT		•					
1999 A							
GQAGLERTSGDPPASASQSAGITGVSHLA*P TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLPFQRLCYY TSMPLLYFQAGGVQWQCDLGSPQPLPFGF TSMPLLYLNS*PIDLPTSPSQSAEIKGVSHR CPASFYLPLKYYLEAKFCA*0ECAPSAGVGA GYKRGHKSCLLINCVVQI GYKRGHKSCLLINCVVQI DAWGPETRLARILNPDSFIEPPGRLPELEATR PHMEPKASCPAAPLMERKFHVLVGVTGSV AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL RRWADLLLVAPLDANTLGKVASGICDNLTC VMRA WDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLINGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLINGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGPRNPE QSFRGAWGPSFWGSWKSQELSAGGAQAWP LLGSAGSGLRGEA CGSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGPRNPE QSFRGAWGPSFWGSWKSQELSAGGAQAWP LLGSAGSGLRGEA CGSPIPSRSSPSCSWAQVPGVLCPPWPFKVLTLQA CGSPIPSRSSSPSCSWAQVPGVLCPPWPFKVLTLQA PFFFFPGGLSLATPGD PAVGLGALGVIPPVVQRPPTQRSQGRGW DPERDPGCRVQVSRGPRPGGGKTGGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS MMNNTTCIQPSMISSMALPITYLLCIVGVFGN TLSQWIFLTKIGKKTSTHTY.SHLVTANLLVC EVHKMMYREFFEERAGASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	649	1999	A	4873	226	189	
TSMPLLPFQRLCVYI	1.						FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV
A	1						
K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG FTMLARLVLNS*PHDLPTSPSQSAEKGVSHR CPASFYLFLKYYLEAKFCA*GECAPSAGVGA GYKRGHKSCLLINCVQ1 651 2001 A 4898 1701 771 DAWGPETRLARILNPDSFIEPRGRLPELEATR PHMEPKASCPAAAPLMEKFHVLVGVTGSV AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL RRWADLLLVAPLDANTLGKVASGICDNLTC VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA SP*IPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPOVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCRLIPSSWDYRHATMPG*FF*YF*R WGFTELAILVLNS*PQVICPPWPPKVLTLQA G54 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EFPEGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHILAASCVVVVCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSCCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSSCS SCSAVSQQEDRCRSSCCS SCSAVSQQEDRCRSSCCS SCSAVSQQEDRCRSSCCS SCSAVSQCEDRCRSSCCS SCSAVSQCEDRCRSSCCS SCSAVSQCEDRCRSSCCS SCSAVSQCEDRCRSSCCS							
	650	2000	A	4874	2	437	
CPASFYLFLKYYLEAKFCA*GECAPSAGVGA GYKRGHKSCLINCVQI							
GYKRGHKSCLLINCVVQI							L
A							
PHMEPKASCPAAAPLMERKFHVLVGVTGSV	651	2001	À	4898	1701	771	
AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL RRWADLLLVAPLDANTLGKVASGICDNLLTC VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQ\$*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA 652 2002 A 4927 I 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPAHTHSRTHPSAPLVKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPFGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIITYLLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD						,,,	
SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL RRWADLLLVAPLDANTIGKVASGICDNLLTC VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA LSLPRGQACSRLNWTEGPGLFSFGGCPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGFRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF***R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPEGLESLATPGD IPAVGLGALGVIPPVRVPQRPTYQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4988 332 159 MNNNTTCIOPSMISSMALPHIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC REVHIKTMR*HFLPIRLEKNKNNIKD							AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY
VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTTVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EFEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHYTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	•						
VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSFWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD							
EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGFGLSFFQPGEAAA 652 2002 A 4927 1 611 FRGRTSRPARGFSPWRPPGTIMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFF1*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEFGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIPLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	1 1						
LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTLAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPITYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD							•
LSLPRGQACSRLNWTEGPGLSFFQPGEAAA 652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAPSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLIKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	l						
652 2002 A 4927 1 611 FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD							
SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	652	2002	A	4927		611	
PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD					-		
GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTLAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPEGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPITYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	1	i				ļ	
KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA 653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTLAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRSWFCQLP*EPPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPITYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD		į			·		
LLGSAGSGLRGEA		ſ	[1	1	ſ	
653 2003 A 4965 2 283 FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD					ļ	l	
VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLIKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	L						
WGFTELAILVLNS*PQVICPPWPPKVLTLQA 654 2004 A 4968 3 437 RPGIPGRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLIKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	653	2003	A	4965	2	283	
654 2004 A 4968 3 437 RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD							
ipavglgalgvippvrvpqrptqrsqgrgw dperdpgcrvqvsrgprfgeqktpglqgclp ppclthlaaascvvvwcgrwkrdsaecqcd hscsavsqqedrcrssscs 655 2005 A 4983 201 397 Minnttciqpsmissmalpiiyillcivgvfgn tlsqwifltkigkktsthiylshlvtanllvc 656 2006 A 4988 332 159 Lvhkdmyreffeeeaqasnkhvtrcctslvi revhiktmr*hflpirleknknnikd	654	2004	- 	1069	2	427	
DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	""	2004	^	4700	<i>3</i> ,	437	
PPCLTHLAAASCVVVWCGRWKRDSAÈCQCD HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD]			j			
HSCSAVSQQEDRCRSSSCS 655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	j i]	- 1	ĺ		•	
655 2005 A 4983 201 397 MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	[[ĺ	ĺ		ĺ		
TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC 656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD	655	2005	A	4983	201	397	
656 2006 A 4988 332 159 LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD						·	
	656	2006	A	4988	332	159	
657 2007 B 5008 129 465 MAGMKTASGDYIDSSWELRVFVGEEDPEAES	اا						
	657	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i	ļ	ľ	·	peptide		/=possible nucleotide deletion, \=possible
<u> </u>		<u> </u>	ļ	sequence		nucleotide insertion
		١	1		Ī	VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
Į.	ľ	l	1			AIWWEQKRQWLLQTHWTLDKYGILADARLF FGPOHRPVILRLPNRRALRLX*
658	2008	A	5017	1	292	
030	2008	^	3017	1	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF
ŀ						HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
659	2009	A	5018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
1007	2007	^	3018	17	338	T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
		1			1	MAACWAVHVKTHMRPGLAVLPRLVLNSWS
		l			ł	*AIILLWPPKALGLQA
660	2010	A	5028	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG
1 555	20.0		3020	_	3.0	HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGST
		l	1			AGSCPROKKTTPGPTVLCVCSFWIYQRGEPH
			·			HRTGARWNH
661	2011	A	5050	752	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
						LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS
1						LELLGSSHPPTSASQSARITGVSHRAWPLK*F
			1			NLNQYQTLTMN
662	2012	Α	5054	48	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH
						EAASQGRLLALRTLLSQGYNVNAVTLDHVTP
						LHEACLGDHVACARTLLEAGANVNAITIDGV
1		i				TPLFNACSQGSPSCAELLLEYGAQAQLESCLP
		1				SPTHEGASKGHHECLDILISWGIDVDQEIPHSG
1						TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
						WDTPLPGAGHQSTQKLE*LFAMVEIWQ
663	2013	Α	5066	951	580	VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK
1		1				ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
						GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
<u></u>						WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
664	2014	Α	5071	550	1 .	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
	•					QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP
1 1			ł			MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK
						TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR
						YMAICNPLRYSVLMGHGVCMGLMAAAWAC
665	2015	Ā	5074	406	692	GFTVSLVTTSLVFHLPFHSSNQHE
003	2013	A	30/4	496	092	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
						PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD
						HDQNEGFHCREECRILGHSDRCWMPRNPMPI
						RSKSPEHVRNIIALSIEATAADVEAYDDCGPT
, ,					·	KRTFATFGKDVSDHPAEERPTLKGKRTVDVT
						ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
∤ ¹				į	ļ	PTKPSVSYEIVDPGITARRC
666	2016	A	5080	408	248	IMLLSTSS*VYFOSSTKDSHFFLFDFOKTGPPL
		· ·	5000			VGPKAQLSGLQLQPCLYKRR
667	2017	Α	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ
		••	"		- ''	PCVY*RR
668	2018	A	5086	852	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
	2010	••	5555	332		RALPTIFADIENLKYLLFIRDASQPFYLGHTV
						IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE
			1	İ		TCADQSVIWKLSEDKQLAICLKYAGVHAENA
		. 1				EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG
]					CCSDMAITFNGLTPQKMEVMMYGLYRLRAF
	i					GHYFNDTLVFLPPVGSEND
669	2019	A	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG
				-		ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP
						RGCQHEAAPCPRGPGSDGLHHASAACASLPP
						SPILPVLLPELGPL
670	2020	Ā	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isolaucine, K=Lysine, L=Laucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ŀ	914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ	ŀ		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1						DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI
1	ł		i		•	VGGTDQQYVSNDSGIYVSRIKENGAAALDGR
						LQEGDKILSVNGQDLKNLLHQDAVDLFRNA
I			1			GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI
			l			FMVLVPVFALTMVAAWAFMRYRQQL
671	2021	Α	5105	672	400	RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF
	1					VLLLLLISLLCLYWKARKLSTLRSNTRKEKA
		}				LWVDLKEAGGVTTNRMED*EEDECN
672	2022	Α	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT
1		}				YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS
					Ì	NQAHGALQEYVLAPCS
673	2023	Α	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA
:	1		1			NHFVEVT
674	2024	A	5153	3	2953	LTEDQPFDILQKSLQEANITEQTLAEEAYLDA
1		1				SIGSSQQFAQAQLHPSSSASFTQASNVSNYSG
1]	ļ]		J	QTLQPIGVTHVPVGASFASNTVGVQHGFMQH
i i						VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS
i						MMTINNLDGSQIILKGSGQQAPSNVSGGLLV
		1]			HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF
	•					QTSLPVHNIIIQRGLAPNSNKVPINIQPKPIQM
						GQQNTYNVNNLGIQQHHVQQGISFASASSPQ
	1					GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG
1						GSIVIHSPMGQPHAPQSQFLIPTSLSVSSNSVH
						HVQTINGQLLQTQPSQLISGQVASEHVMLNR
						NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ
						TFAASGSPVIANHASPQLVGGQMPLQQASPT
1						VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR
	•		1			FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL
1	!					TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT
1		i	1			GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR
}						QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ
1	· '					QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ
1 .						LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD
						AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV
						ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE
						MVMIDRMFNQEERASLSRDKRLALVDPEGFQ
						ADFCCSFKLDKAAHETQFGRSDQHGSKASSS
						LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP
] :		ļ .]	NHIVVSAEGNISKKTECLGRALKFDKVGLVQ
						YQSTSEEKASRREPLKASQCSPGPEGHRKTSS
		· ·				RSDHGTESKLSSILADSHLEMTCNNSFQDKSL
		!				RNSPKNEVLHTDIMKGSGEPQPDLQLTKSLET
						TFKNILELKKAGRQPQSDPTVSGSVELDFPNF
						SPMASQENCLEKFIPDHSEGVVETDSILEAAV
			J			NSILEC
675	2025	Α	5154	599	1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR
						LYDEVQEVVYFPAVVHDNLGERLKCTYIEID
						QVPETYAVVLSRPAWLWGAEMGANEHGVCI
		.				GNEAVWGREEVCDEEALLGMDLVRLGLERA
						DTAEKALNVIVDLLEKYGQGGNCTEGRMVF
						SYHNSFLIADRNEAWILETAGKYWAAEKVQE
						GVRNISNQLSITTKIAREHPDMRNYAKRKGW
						WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE
					1	GYKLLNKHKGNITFETMMEILRDKPSGINME
						GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER
						SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS
						HFKPDRRHPLYQKHQQALEVVNNNEEKAKI
						MLDNMRKLEKELFREMESILQNKHLDVEKIV
L						NLFPQCTKDEIQIYQSNLSVKVSS
						

No. of a	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine G=Glycine, H=Histidine
Sequence							I=Isoleucine, K=Lysine, I =I encine
1914 1915 1916	1						
amino acid residue of peptide residue of peptide sequence T-Threonine, V-Valine, W-Tyrotine, N-Stop codon, peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide sequence Peptide servino Peptide sequence Peptide servino P							
Peptide Sequence						of peptide	T=Threonine, V=Valine, W=Tryptophan,
Peptide		İ			residue of		
	ļ.	}	l	ŀ	peptide		/=possible nucleotide deletion, \=possible
FTPESCLSLPSSWDYRRPPRRAMIT_YF**RRG FTHLARWSIS*PIDPPASASQSAGITGVSHRA RT FTHSYDLLALEOSKTTYKPDWFDIVESEVGC FTHSYDLLALEOSKTTYKPDWFDIVESEVGC KEAVCVIDMSSTEFEITSTGDQALEVLQYLF SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR LYRSFMISTDQCVICWAWLKKIMRKDS NLLLEDVTWKYTALNLIGRAVDVLSELSYA PMTPDHFTSLPCKEMSVGYANGIRWASMTHT GEPGFMLYTPEYRWGFTMLSTLVSNS GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVYTSKSDWWAFG GRIAKMPYKWIAESLARRVXTAVTLLARD GRIAWWAFALYKWIAESLARRVXAAVTLLARG LLICLSYLFILSAIRUYSAAGRCKAFSTCSAH RTVVVVYGTISMYFKPKAKDPNVDKTVAL FYGVTPSLNPIDINSINAMAVKAVILLAGG LLISKASHCYCCPLPLSAGIG GRIAWWAFALYKWIAESLARRVXAAVTLLAGG LLISKASHCYCCPLPLSAGIG GRIAWWAFALYKWIAESLARRVXAAVTLAGG LLISKASHCYCCPLPLSAGIG GRIAWWAFALYKWIAESLARRVXAAVTLAGG GRIAWWAFALYKWIAESLARRVXAAVTLAGG GRIAWWAFALYKWIAESLARRVXAAVTLAGG GRIAWWAFALYKWIAESLARRVXAAVTLAGG GRIAWWAFALYKWIAESLARRVXAAVTLAGG GRIAWWAFALYKWIAESLARRVXAAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTKGGT GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFALAKWATAVTLAGG GRIAWWAFATAKGGG GRIAWWAFATAKGGG GRIAWWAFATAKGGG GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAKGGT GRIAWWAFATAK					sequence		
FTILLARNVSIS*PHDPPASASQSAGITGVSHRA RT	676	2026	Α	5155	2	306	FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG
RPT	1		ŀ	ľ		<u> </u>	FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG
	i			İ			FTLLARMVSIS*PHDPPASASQSAGITGVSHRA
			L				
SNDLDVPVGHIVHTGMLNEGGYPENDCSIA	677	2027	A	5167	97	740	
Linkrsfmisptdoguhcwawlikihmpkds			1				
						1	
PMTPDHPSLFCKEMSVGYANGRIVMSMTHT			1			ł	
GERGENLYPIETYRUGFILSTLYNSS	1 :		ľ				
2028							
GRIAKMPYKWIAIESLADRYYTEKSDYWAFO							
	678	2028	A	5183	1919	2018	
HRILKOPEDCLIDELCKI**SPOSP							
679 2029 A 5190 39 499 RESQVKHFKMRKIDLCLSSEGSEVILATSSDE KHPPENIDGNPETFWTTTGMFPQEFIICFHKH VRIERLVIQSYFVQILKIEKSTSKEPVDFEQWI EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS AFDHFASVHSVASGTVSNLSS SPDHFASVHSVASGTVSNLSS ELLAVLKLACGGISLNALALMVATAVLTLAPL LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH RTVVVVFYGTISRMYKFPKAKADFNVDKTVAL FYGVVTPSLMFIIVSLANGEVKAAVLTLLRGG LLSRKASHCVCCPLPLSAGIG VPDNGDVTKLPVCSTLVEETSLTVSEAMEQS KNESPLPGTLAHTCNTSTLGGGRGWIT*GREF DISMAMMVRPCLYRK SELVENDEN STATEMEN S							
KHPPENIIDGNPETEWTTTCMEPQEFIICFHKH		0000		44.00			
	פ/ס	2029	A	2170	39	499	
EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS							
AFDHFASYHSVSAEGTVVSNLSS							
LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH	400	2020		5204	541	02	
RTVVVVPGTISFMYFKPKAKDPNDKTVAL FYGVVTPSLNPIPYSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMYKPCLYRK SELSPLPSSTROAGVQWPNLSSLKTLPPGFK*P SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WILLESMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LF*1PADDSPTDQEPKTVMLSK QSESA1 WLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYKFSATLVKIPQHL TDS TDS WRTTPLF*1PADDSPTDQEPKTVMLSK QSESA1 WLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYKFSATLVKIPQHL TDS WRTTPLF*1PADDSPTDQEPKTVMLSK QSESA1 WLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYKFSATLVKIPQHL TDS WRTTPLF*1PADDSPTDQEPKTVMLSK QSESA1 WLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYKFSATLVKIPQHL TDS WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLSK WRTTPLF*1PADDSPTDQEPKTVMLST WRTTPLF*1PADDSPTDQEPKTVMLSEKQFFAEEW ORKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMGKLSCT WWQLAQPNLGTILDPRPGVTIKGFTQLPGD AIYHISDLEDEEEGITFQVQQPLEVEEKLSTS WPVTGIFLPTTSAGGPVTVATANPGKCLSCT WWQQLAQPNLGTILDPRPGVTIKGFTQLPGD AIYHISDLEDEEEGITFQVQQPLEVEEKLSTS WPVTGIFLPTTSAGGPVTVATANPGKCLSCT WSTTFTTCRLHFSDITQVTPSSGFPSLSCGSS	000	2030	A	3204	341	92	
FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK DTSMANMVKPCLYRK SCLSLPSWDYRCLPPCPANFCIPSRNGVLPC WPGWSTPDLS SCLSLPSSWDYRCLPPCPANFCIPSRNGVLPC WPGWSTPDLS							
681 2031 A 5207 10 247 VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCHTSTLGGRGRWIT*GREF DTSMANMVRPCLYRK 682 2032 A 5210 2 231 FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRIPDLS 683 2033 A 5218 85 402 CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK QSESAI 684 2034 A 5220 1 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YC*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRREGGGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKHLIGVLSIVEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSPEGGILLFILTKR EVDKDRVKQMKAQNMRLSNTGEYESQRFR ASSQSAPSPDVGSCVTPTESLASLCTTQS ETDLSSASCLTQSSSEVAGSSQKMGQFGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQREGVSGCVTPTESLASLCTTQS ETDLSSASCLERGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPQVITKGFTQLPGD ATYHISDLEEDEEGGITFQVQQLEVEEKLSTS KPYTGIFLPPITSAGGGPVTVATANPGKCLSCT NSTFIFFITCRILHPSDITQV	1		i				1
KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK	681	2031	Ā	5207	10	247	
DTSMANMVKPCLYRK	***	2031	**	3207	10	24,	
682 2032 A 5210 2 231 FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS 683 2033 A 5218 85 402 CPSVSGLKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFI*PADSPTDQEPKTVMLSK QSESAI 684 2034 A 5220 1 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSULIHHR KHSRPIVTWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLIGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSASPSPDVGSCVQT 688 2038 A 5249 1 1407 LQQTEDKSLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKJQVLADQKEGVSGCVTPTESLASLCTTQS ETITLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVTIKGFTQLPGD AIYHISDLEEDEEGGTTFQVQQPLEVEEKLSTS KPYTGIEIPTTSAGGPVTVATANPGKCLSCT NSTFIFTTCRILHPSDITQVTPSSGFPSLSCGSS	ŀ						
SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS	682	2032	Α	5210	2	231	
WPGWSRTPDLS					-		
MIILLRSMYRINVKPYFF1*LFFSRVNC*SVIIG	1 1						
YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK QSESA1	683	2033	A	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF
QSESAI							MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG
684 2034 A 5220 1 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*\$*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPELSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK
YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEFEPLSPELEYI PRKRGKNPMKAVGLA WAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							QSESAI
YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS	684	2034	Α	5220	1	194	
685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL
OKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
686 2036 A 5239 79 508 GGEAAARAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	685	2035	Α	5228	260	440	
AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	<u> </u>						
KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	686	2036	A	5239	79	508	
VIQNSKRKGPETIKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	[' [
687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVTTKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	<u></u>						
PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	687	2037	A	5244	1	428	
EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS			- 1				•
ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					ļ		
688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS		İ					
SGDSDLATALHRLSLRRQNYLSEKOFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	688	2038	A	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP
EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS			1]		SGDSDLATALHRLSLRRQNYLSEKQFFAEEW
HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS			İ		ļ	1	
AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS			ŀ				EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY
KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS		•		,	l	l	HWQQLAQPNLGTILDPRPGVITKGFTQLPGD
NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS				'		J	
	1	1		Ī			
GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT	.		. [ļ	ŀ	
	L						GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

SEQ II	D SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of		hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uenœ		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	J	j]	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	- 1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	ł	1	peptide	ł	/-possible nucleotide deletion, \-possible
 		├	 	sequence	ļ	nucleotide insertion
1		ļ	}	l .		FSSTMSLAKLLQERGISAKVYHSPISENPLQPL
	1	İ				PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN
			1			FLASRPAETFLQEMYGLRPSRNPPDVGQLKM
1		ł	1	l	1	NLVDRLKRLGIARVVKNPGAQENGRCQEAEI
	1				J	GPQKPDSAVYLNSGSSLLGGLRRNQSLPVIM
689	2039	A	5254	2	2621	GSFAAPVCTSSPKMGVLKED
1 005	1 2035	1	3234	1 ~	2021	LSLFGSRALGRSGARAMAKAKKVGARRKAS
	ł	1		}		GAPAGARGGPAKANSNPFEVKVNRQKFQILG RKTRHDVGLPGVSRARALRKRTQTLLKEYKE
	1	1]		RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA
	1					LEQQRHHEKKSIYNLNEDEELTHYGQSLADIE
1	ſ	ĺ			[KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG
						GLLHKKTQQEGEREKPKSRKELIEELIAKSK
1		Į			1	QEKRERQAQREDALELTEKLDQDWKEIQTLL
	ŀ					SHKTPKSENRDKKEKPKPDAYDMMVRELGF
		1			ſ	EMKAQPSNRMKTEAELAKEEQEHLRKLEAE
	1	ł	1			RLRRMLGKDEDENVKKPKHMSADDLNDGFV
ļ	1		1		}	LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA
}						SDPESNEEEGDSSGGEDTEESDSPDSHLDLES
1						NVESEEENEKPAKEQRQTPGKGLISGKERAG
1	1		·			KATRDELPYTFAAPESYEELRSLLLGRSMEEQ
						LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE
						YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP
[ESASDAIKFVLRDAMHEMEEMIETKGRAALP
1			1			GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL
} .	i l					SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS
ļ						QRFIPELINFLLGILYIATPNKASQGSTLVHPFR
						ALGKNSELLVVSAREDVATWQQSSLSLRWA
İ	1 1		ĺĺ	·	·	SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
l	1 1			ł		YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ
						ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK
						REFKGAVREIRKDNQFLARMQLSEIMERDAE
1.	1 1		1 1			RKRKVKQLFNSLATQEGEWKALKRKKFKK
690	2040	A	5261	1 :	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW
	j j]]]	ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL
l	1					FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT
Ĺ			1		1	SFVK
691	2041	A	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
						EVLSSFFFFFLKFSYKPONIV
692	2042	A	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV
			[Į.		ERVLTFLPAKALLRVACVCRLWRECVRRVLR
	1			ł	ł	THRSVTWISAGLAEAGHLEGHCLVRVVAEEL
					ļ	ENVRILPHTVLYMADSETFISLEECRGHKRAR
		j		ļ	j	KRTSMETALALEKLFPKQCQVLGIVTPGIVVT
İ		ł		- 1	.[PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF
I		1		ł	ı	HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV
ı		}	. 1	ĺ	i	FGYNCCKVGASNYLQQVVSTFSDMNIILAGG
		- 1	į	ľ	l	QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
		1	ļ		İ	QSATVLLNEDVSDEKTAEAAMQRLKAANIPE
	1	1	1	{	į	HNTIGFMFACVGRGFQYYRAKGNVEADAFR
]		ļ	j	KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
<u> </u>	10040	<u> </u>				EVKDDDLFHSYTTIMALIHLGSSK
693	2043	A .	5301	362	507	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
694	12044		6216		-	ACFPTNIVTLCHSIA
J74	2044	A	5310	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK
•		İ		j		KTRAMRRLNMHEENLKTKKQHRKERLYPL
695	. 1	- {	1	-1	1	RKYAAKA
	2045	A	5315	125		ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
Į			1			TGSGKEEGPAPCKQMKLEAAGGPSALNFDSP SSLFESLISPIKTETFFKEFWEQKPLLIQRDDPA
		İ				LATYYGSLFKLTDLKSLCSRGMYYGRDVNV
						CRCVNGKKKVLNKDGKAHFLQLRKDFDQKR
•						ATIQFHQPQRFKDELWRIQEKLECYFGSLVGS
						NVYTTPAGSOGLPPHYDDVEVFILOLEGEKH
						WRLYHPTVPLAREYSVEAEERIGRPVHEFML
						KPGDLLYFPRGTIHQADTPAGLAHSTHVTIST
1						YONNSWGDFLLDTISGLVFDTAKEDVELRTG
1						IPRQLLLQVESTTVATRRLSGFLRTLADRLEG
						TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP
						GGKLPRLDSVVRLQFKDHIVLTVLPDQDQSD
ſ						ETQEKMVYIYHSLKNSRETHMMGNEEETEFH
						GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE
						EKESLVLSLWTECLIQVV
696	2046	Α	5318	1476	742	LMKXYLEAAELGEISDIHTKLLRLSSSQGTIET
						SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN
1						VEKLQVLLNCMTEIYYQFKKDKAERRLAYN FEOTIFYEDROVI VYLATIA ATTUTTEDE OVEK
						EEQIHKFDKQKLYYHATKAMTHFTDECVKK
						YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI EEEVSKYQEYTNELQETLPQKMFTASSGIKHT
						MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE
						LAENNHILESGGSLTMDGGLRNVDCL
697	2047	Α	5320	244	478	LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP
"	20		0020		.,,	PPGFKOFSCLSLPSSWDYRHLPPHLANFSREG
()						VSPSWPGWSRTPDFR
698	2048	Α	5324	266	714	LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC
1						LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP
1						VGGLLMAFQKYSGETVQERKQKDRKALHEL
						KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA
				32		KKIEALLNLPRNPSVIDKQDKD
699	2049	A	5334	699	277	RPHGHLVCISSSAGLSGVNGLADYCASKFAA
					:	FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM
1						FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM
						YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK
700	2050	Α	5344	3	614	PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE
1 /00	2030	^	2277	,	014	QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS
						QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
1						PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
						VETIQAQLLSTHDQPSVQALADEKNGAQTRP
						AGAPPRSLOPPPGRPDPELSSSPRRGPPKDKK
				•		LLATNGTPL
701	2051	Α	5346	3	1383	HASVLFCRVMAASKTQGAVARMQEDRDGSC
					•	STVGGVGYGDSKDCILEPLSLPESPGGTTTLE
				ļ		GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV
						IADVKLVADFQRYILYWRKRFTEQPITDFCSV
1						IRINSTAPFEEQENYFLLCDVLPEDRILREELQ
]]			j	J		KQRLREILEQQQQERNDTNFHGVCMFCNEEF
		1		l	ļ	LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL
]			1	İ		CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR
				l		KKQHRKINPKNREYDRFYVINYLELGKSWEE
		ļ				VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKOAETIEKLYVHMEDAHEFDLLKIKSELG
		l		!		LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS
				I	·	KADLRTHMEETKHTSLLPDRKTWDQLEYYFP
				ł	ŀ	TYENDTLLWTLSDSESDLTAQEQNENVPIISE
				ļ		DTSKLYALKQSSILNQLLL
702	2052	A	5356	2502	1540	MAAATRGCRPWGSLLGLUSAAAAAWD
		لــــــــــــــــــــــــــــــــــــــ				

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	согтеspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
neuce		ł	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
		L		sequence		nucleotide insertion
						LASLRCTLGAFCECDFRPDLPGLECDLAQHL
					•	AGQHLAKALVVKALKAFVRDPAPTKPLVLSL
						HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH
1						HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA
]			1			CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV
						VYGTNYRKAIFIFISNTGGEQINQVALEAWRS
			i			RRDREEILLQELEPVISRAVLDNPHHGFSNSGI
1		ł	,			MEERLLDAVVPFLPLQRHHVRHCVLNELAQL
					·	GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK
703	2053	A	5380	270	657	TVASRIAFFL
1 /03	2055	Α.	2200	278	037	LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL EERLEFWMEKYDKDTEMKQNELNALKATKA
ļ i]		Į	SDLAHLODLAKMIREYEOVIIEDRIEKERSKK
]			1			KVKQDLLELKSVIKLQAWWRGTMIRREIGGF
704	2054	A	5381	1	1003	KM FRGRAVKMAAVVEVEVGGGAAGERELDEV
/ ~	2054	^	3361	•	1003	DMSDLSPEEQWRVEHARMHAKHRGHEAMH
1 1						AEMVLILIATLVVAQLLLVQWKQRHPRSYN
1						MVTLFQMWVVPLYFTVKLHWWRFLVIWILF
1 1			}		İ	SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY
.						KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA
						MDFGISLLFYGLYYGVLERDFAEMCADYMA
						STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV
1 1						SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK
1						QTCPYCKEKVDLKRMFSNPWERPHVMYGOL
i i			1			LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	A	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP
						GNKRPLSALYRLESKEPFLSVGGYVFDYDYY
					٠	RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA
[[•					VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS
						DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK
i i						AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP
l i						AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ
1						IK
706	2056	A	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI
						S
707	2057	Α	5415	6	287	PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT
						VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP
						EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	Α	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS
						TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS
<u> </u>						GCTIFKTVTLTARSTASLLKSVRPRTHQKE
709	2059	Α	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG
						SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ
	·					VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK
710	2000			1070		GKWFCPRCVQEKRKKK
710	2060	Α	5442	1073	559	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS
]					!	GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY
1 [ĺ			[,	EESDSDESWTTESAISSEAILSSMCMNGGEEK
		l		ļ		PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI
						RVRKPFKCRCGKSYKTAQGLRHHTINFHPPV
71.	2061		5440		210	SAEIIRKMQQ
711	2061	A	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP
		ļ			ļ	DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL
1		1				NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD
712	2062	_	5400	<u> </u>	740	LYRDIPELQGF
/12	2062	A	5499	91	749 .	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL
	ł	- 1	1		ļ	QVRGEALSEEFIWSLLFLAAEQLLEDLRNDSS
L		1				DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

COCO TO	L CTO ID	1 17.3	Lara	10-11-11	FB - 2'-7-13	[A
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		ino	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]	J	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			į.	peptide	3040200	/=possible nucleotide deletion, \=possible
		l	į.		l	nucleotide insertion
		ļ	ļ	sequence		
1		1.	1	1	l	KAPELLQGQSEDEQPDASQMHVYSLGMTLY
1	ĺ	1	1	1		WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH
1		1	1			RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR
1	}	1	1	l		RLVGLVLGTISEVSREPCFSSSSCWSCVAIKI
713	2063	A	5506	22	478	VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT
į.	İ	1	!	ł	ł	TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL
ŀ			1		İ	GGVECLLLAVMAYDRCVAICKPLHYMVIMN
1		ļ	j			PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR
		1	ł			
	0051				 	CGHHEVDHFLCEMPALIRMACISTV
714	2064	Α	5514	25	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR
J]	J	}			RLTSSVSCALDEAAAALTRMRAESTANAGQS
	ł		1			DK
715	2065	A	5526	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM
1		1			ŀ	GRTALFHHSGGSSGYESLRRDSEATGSASSAP
1					l	DSMSESGAASPGARTRSLKSPKKRATGLORR
1	ł	ì	1	ľ	1	RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG
1	i	I	1	ļ		SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG
1	1				•	
					•	LACVSTRLRLAERRQQRLREVQAKHKHLCEE
1	1				1	LAETQGRLMLEPGRWLEQFEVDPELEPESAE
1			1		ĺ	YLAALERATAALEQCVNLCKAHVMMVTCFD
L		l				ISVAASAAIPGPQEVDV
716	2066	A	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD
						SSSSERTMDLVLEMCNTNSIHWCGISGRQLG
i			i			KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF
			i			LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG
'1'	2007	A	2331	3	400	
1		1	l			RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS
1						VNIREELKLANVVFFPRCLLVQRCGGNCGCG
1			,			TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR
	l					RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2068	A	5586	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDF
i	1		1			TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS
	1		1			KAEEAAQENA
719	2069	A	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM
1 '''	2007	١.,	3370	•	330	LATNSTRGLNEDELMAHGQEKDSSSESEDSC
				1		PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL
L	20==	<u> </u>			·	FPLAFGLFNIVAAERC
720	2070	Α	5628	798	148	LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP
	}	ł		·		LLLRCKLLPSALQKMALGMFFGFTSVIVAGV
		l				LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW
						WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG
					-	AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH
					·	CPKDFGNINNCRMDLYFFLLAGIQAVTALLF
[[
701	2071		5620	146	606	VWIAGRYERASQGPASHSRFSRDRG
721	2071	Α	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA
1]			KLHETALHHAAKVKNVDLIEMLIEFGGNIYA
						RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT
						LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID
1						YLSYN
722	2072	A	5638	3	3806	CPSLDIRSEVAELROLENCSVVEGHLOILLMF
				-		TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES
						LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD
[VALPALGAVLRGAVRVEKNQELCHLSTIDW
					İ	GLLQPAPGANHIVGNKLGEECADVCPGVLGA
						AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP
					'	HGMACTARGECCHTECLGGCSQPEDPRACV
'						ACRHLYFQGACLWACPPGTYQYESWRCVTA
						ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT
1 1			1		ŀ	
<u> </u>			L			RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	ucaice	İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
-				peptide	sequence	/=possible nucleotide deletion, \=possible
Í	1	ľ				nucleotide insertion
<u> </u>		<u> </u>		sequence		L
		i				QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL
Ì				·		GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD
ł						AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI
				ł		PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN
į.	Į.					KAEINPRTNGDRAACQTRTLRFVSNVTEADRI
ì						LLRWERYEPLEARDLLSFIVYYKESPFQNATE
ŀ	İ	Į.				HVGPDACGTQSWNLLDVELPLSRTQEPGVTL
ł	ł	i	-	1	Ì	ASLKPWTQYAVFVRAITLTTEEDSPHQGAQS-
						PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW
		ŀ				KPPTQRNGNLTYYLVLWQRLAEDGDLYLND
ł						YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD
	ļ		1			CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH
	1			i	·	NAITIPISPWKVTSINKSPORDSGRHRRAAGPL
}		1				RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY
		1				RIDIHACNHAAHTVGCSAATFVFARTMPHRE
						ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL
Į	ł	l	ļ	1	}	ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV
ŀ		ŀ	ł			HLALLPPGNYSARVRATSLAGNGSWTDSVAF
ĺ			1			YILGPEEEDAGGLHVLLTATPVGLTLLIVLAA
]				I .
	ŀ					LGFFYGKKRNRTLYASVNPEYFSASDMYVPD
	İ			į		EWEVPREQISIRELGQGSFGMVYEGLARGLE
						AGEESTPVALKTVNELASPRECIEFLKEASVM
					İ	KAFKCHHVVRLLGVVSQGQPTLVIMELMTR
ł			İ	ł	1	GDLKSHLRSLRPEAENNPGLPQPALGEMIQM
1						AGEIADGMAYLAANKFVHRDLAARNCMVSQ
ĺ						DFTVKIGDFGMTRDVYETDYYRKGGKGLLP
		i				VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV
	ļ-			1		TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC
						PLQLQELMSRCWQPNPRLRPSFTHILDSIQEEL
						RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP
1						TPRDCSPQNGGPGH
723	2073	A	5672	1	216	LAWIDNILPEKEKKETDKKRKKKKAHEDCD
	1		,			EEPOFPPPSVIKIPMESVQSDPQNGIHCIARKR
i						SSSWSYSL
724	2074	A	5704	4235	940	ARGRRSRPVWAASWGGRGRPAARRRPRGLA
'-]	• • • • • • • • • • • • • • • • • • •	1	- 1	ATMGFELDRFDGDVDPDLKCALCHKVLEDP
		1				LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR
	1	l				LSAKELNHVLPLKRLILKLDIKCAYATRGCGR
		1				VVKLQQLPEHLERCDFAPARCRHAGCGQVLL
			· .			RRDVEAHMRDACDARPVGRCQEGCGLPLTH
		1				GEORAGGHCCARALRAHNGALQARLGALHK
1	1					
		1				ALKKEALRAGKREKSLVAQLAAAQLELQMT
1		1		1		ALRYQKKFTEYSARLDSLSRCVAAPPGGKGE
1				Í		ETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDG
		1		ľ		SSSEGIFVSKIVDSGPAAKEGGLQIHDRIIEVN
		1				GRDLSRATHDQAVEAFKTAKEPIVVQVLRRT
		1				PRTKMFTPPSESQLVDTGTQTDITFEHIMALT
1		1				KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI
		ŀ				GDIHQEMDREELELEEVDLYRMNSQDKLGLT
]		l				VCYRTDDEDDIGIYISEIDPNSIAAKDGRIREG
		l	}			DRIIQINGIEVQNREEAVALLTSEENKNFSLLI
1	l	1	}	ł		ARAELQLDEGWMDDDRNDFLDDLHMDMLE
	1				,	EOHHQAMQFTASVLQQKKHDEDGGTTDTAT
						ILSNOHEKDSGVGRTDESTRNDESSEQENNG
1						DDATASSNPLAGQRKLTCSQDTLGSGDLPFS
1 .						NESFISADCTDADYLGIPVDECERFRELLELK
						COVKSATPYGLYYPSGPLDAGKSDPESVDKE
1		l				LELLNEELRSIELECLSIVRAHKMOOLKEOYR
						ESWMLHNSGFRNYNTSIDVRRHELSDITELPE
1	1	l				
		ı	ı	1	ĺ	KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRR

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide			USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
,	seq-		1		corresponding	1
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ì	ļ	J	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł				peptide		/=possible nucleotide deletion, \=possible
	L		L	sequence		nucleotide insertion
				1		AAEGISCPSSEGAVGTTEAYGPASKNLLSITE
Į.	Į.	1	ł		İ	DPEVGTPTYSPSLKELDPNQPLESKERRASDG
	l		1		1	SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA
	İ		1			QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK
			1			DLSSPTPSEPRMEWKVKIRSDGTRYTTKRPVR
1	1	ł	1	Ì	ł	DRLLRERALKIREERSGMTTDDDAVSEMKM
					Ī	GRYWSKEERKQHLVKAKEQRRRREFMMQSR
1					•	LDCLKEQQAADDRKEMNILELSHKKMMKKR
ŀ	j	1	İ	İ		NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF
	1	l	1			1
70.5	2025	 	6505		1220	LSVTTV
725	2075	A	5707	3	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP
		!	1			DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY
1		l	1		1	LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG
1	1	1	1			QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW
	1	1				LNIYYIVIISWAIYYLYNSFTTTLPWKQCDNP
	-	1				WNTDRCFSNYSMVNTTNMTSAVVEFWERN
						MHQMTDGLDKPGQIRWPLAITLAIAWILVYF
		İ				CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV
				l		TLPGAKEGILFYITPNFRKLSDSEVWLDAATO
			i		l	IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC
ł				ļ	ļ	CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV
						AASGPGLAFLAYPEAVTQLPISPLWAILFFSM
i						-
l .						LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR
1						ELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYS
1						ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE
1	ł	ŀ				MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ
ľ		ł				MTPLTMGNYVFPKWGQGVGWLMALSSMVL
· ·		ŀ	i	ļ		IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV
ĺ		l	L			RPENGPEQPQAGSSTSKEAYI
726	2076	Α	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA
						PONTFLGTIIRKFEGONKKFIIANARVONCAII
						YCNDGFCEMTGFSRPDVMQKPCTCD
727	2077	A	5716	3	274	HASEYFFKLCSFOVFLSFPLATIVIDVGLVVIP
''		l	• •	-	1	LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL
	}					AWFEKMTCYLOLLFNICLPDVSEE
728	2078	A	5737	1899	649	IOASRASPYPRVKVDFALSCHEDLLAPISEPIE
/20	20/0	^	3/3/	1077	U47	
1		1	[WKYHSPEEEISLGPACWLWDFLRRSQQAGFL
					,	LPLSGGVDSAATACLIYSMCCQVCEAVRSGN
						EEVLADVRTTVNQISYTPQDPRDLCGRILTTC
			[YMASKNSSQETCTRARELAQQIGSHHISLNID
1						PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL
]						ALQNVQARIRMVLAYLFAQLSLWSRGVHGG
						LLVLGSANVDESLLGYLTKYDCSSADINPIGG
]					ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE
						LEPLADGQVSQTDEEDMGMTYAELSVYGKL
						RKVAKMGPYSMFCKLLGMWRHICTPROVAD
						KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE
						DNRFDLRPFLYNTSWPWOFRCIENOVLOLER
						AEPOSLDGVD
729	2079	A	5741	1	5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP
127	20/3	^	3/41		טו לנ	
1						PASRRLRAPGSRPRLAPCTRRAAQPAHARMA
						PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL
						LLLLLGAARAGALEIQRRFPSPTPTNNFALDG
						AAGTVYLAAVNRLYQLSGANLSLEAEAAVG
-						PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL
'						QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV
						AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS
'						TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF
						PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT

SEQ ID NO. of nucleotide peptide coiled and peptide sequence (A-Allanine C-Cysteline, peptide coiled and peptide sequence (A-Allanine C-Cysteline, peptide s			1				
peptide cotide to corresponding to corresponding to the stamino and residue of peptide sequence propriet sequence of peptide sequence propriet sequence prop	SEQ ID	SEQID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Docation Docation			poq	•	, , ,	b	
pence unice of the corresponding on the state of the control of th			1			1	
September Sept	1	1 -		1			
mino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence sequence peptide se	1 *	ucance .		1			
residue of poptible sequence Poptible unclostic decision, popsible moleotide discription Popsible unclostic decision, popsible moleotide discription Popsible unclostic decision, popsible moleotide insertion Popsible unclostic decision, popsible moleotide insertion Popsible unclostic decision, popsible moleotide insertion Popsible unclostic decision Popsible unclosed	uence		i	914			
peptide sequence sequ	İ	ł					
moleotide insertion FDINFSDDMILKING GAKEQHILGFYSAFLIP SDPPFGAQSYAYLALINSEARAGDKESQARSL LARICLPHOAGGGDAKKLITSTYJQLG,QCAGG AGRGDLYSRLYSVPPARERLFAVFERPQOSPA ARAPAALCAFRAPAKALTSTYGLGLQCAGG AGRGDLYSRLYSVPPARERLFAVFERPQOSPA ARAPAALCAFRAPAKALTSTYGLGLQCAGG AGRGDLYSRLYSVPPARERLFAVFERPQOSPA ARAPAALCAFRAPAKALTSTYGLGLQCAGG AGRGDLYSRLYSVPPARERLFAVFERPQOSPA ARAPAALCAFRAPACAFAGTSV AVASVNNYTAVELGYNKARICLYBEATGCYPG APDVVAVLDSVVQGTGPACTEKILNIQLQPEQ AVASVNNYTAVFLGTYNKRALIKINLINESMQ VSRRVVTVAYGEPVHHVMQDPADSGYLY LMTSHQMARVKVAACNYHSTGCDCVGAAD AYCGWCALFTRCTLQDCTINSQOJHWTNA SEGFSRCPAMTVLPSEDVRQCPYPGMILQISGS LPSLSGMEMACDYGNIRTVARVPOPAFGHQ IAYCNLLPRDQPPPFPNQDHVTVEMSVRVN GRNIVKANFTTYDCSSTAQVYPHTACTSCLSA QWPCTWCSQQHSCVSNQSRCEASPHYTSVQD CPRTLISPLAPVPTGGSQNILVPLANTAFFQQ AALECSFGLEEIFEAVWVNESVVRCDQVVLH TITRSQVPFLSQLKGPARPLDSPPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RRGPLQPMAGTCPAPFRAFEHSGPLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFEHSGPLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFEHSGPLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFEHSGPLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFEHSGPLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFEHSGPLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFEHSGPLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFEHSGVAGGTRITIT ACMMENTALTHAPPTACHTAPGENGAGTRITIT ACMMENTALTHAPGTAPGAPGLOGT LLTRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSERIVCVTGPAFRAFTHAPGTSGCAGTTVTASKE GRSRLSTALDAPTTACAGGTRATTAPGTACHT WYMQNVSMAVHHIGREFTLCVUNSTLTCTESP GALSNASAVDFTRITITLGGSETAITATSVTCSW LLLLSVVALFYFCKSRRAERYWQKTLLQME EMESORGERKGGAELQTOMTDLTKELBAGCGSTBLC GSLESSLLDNKHFLIVFYHALELQUKDFAVKSTL GSLESSLLDNKHFLIVFYHALELQUKDFAVKSTL GSLESSLLDNKHFLIVFYHALELQUKDFAVKSTL GSLESSLLDNKHFLIVFYHALELQUKDFAVKSTLAG SCLASSLLTALHGRUCLAKYQGNKGSDATIG KARYTLNEEWLLRENGEARPRNLNVSFQCGC MSLSSVRAMDTDTLACAKGCMSDATG KARYTLNEEWLLRENGEARPRNLNVSFQCGC MSLSSVRAMDTDTLACAKGCMSDATG KARYTLAGER GSLESSLLDNKHFLIVFYHALLELGWKPESCHAM ROMITORYTSCVSEA ROFFREFWANLANTSTYGAK ROFFRAMTARTOLQUKFPQVDLM TIDHDACLSVLAGAFRNLELQMG VANGKGNQRNMGSPQPSL	İ				1	sequence	
FOLMPSDONILKKQGAKEQHRLGFYSARLIB SPPPGAQSYAVLALINEARAGNESQARSI, LARICLPHGAGGYAKULIBSYIQLGLQCAGG AGRGDLYSRLVSVPARERIAAVERPQOSPA ARAAPALCAFRADVRAAIRAARTACFVEP APDVVAVLDSVVQGTPACERKLINIQLQPEQ LDCGAAHLQHPLSILQPLKATFVFRAPGLTSV AVASVAVLDSVVQGTPACTEKLINIQLQPEQ LDCGAAHLQHPLSILQPLKATFVFRAPGLTSV AVASVAVLDSVVQGTPACTEKLINIQLQPEQ LDCGAAHLQHPLSILQPLKATFVFRAPGLTSV AVASVAVLDSVVQGTVANGTPAGHAGAGV VVSRRVVTVAYGEPVHHVMQDPADGGVLY LAMTSHQMARVEXACVINISTICOLOCYGAAD AYCGWCALETRCTLQDCCTNSSQQHFWTTSA SGISRCFAMTVLPSEDVRQEPYFGMLQISCS LPSISGMEMACDVGNNIRTVARVPGPAFGHQ AYCOLLPDQPPPFPPPPQDHVTVENSVEVN GRNIVKANFTTYDCSSTAQVYPHTACTSCLSA QWPCTWCSQQHSCVGNSGRCASPPTSPDD CFRILLSPLAPVPTGGSQNILVPLANTAFFQG ALBECSGGLEBGLGHLCMWSDCC RLRGPLQPMAGTCPAPERLABPLSCPPATTVA TYNKAMGSPDCSGCGREDLGHLCMWSDCC RLRGPLQPMAGTCPAPERLABPLSCPPATTVA TYNKAMGSPDCSGCGREDLGHLCMWSDCC RLRGPLQPMAGTCPAPERLABPLSCPPATTVA TYNKAMSPPTASCGGREDLGHLCMWSDCC RLRGRLQPMAGTCPAPERLABPLSCPPATTVA TYNKAMSPPTASCGGREDLGHLCMWSDCC RLRGRLQPMAGTCPAPERLABPLSCPPATTVA TYNKAMSPPTASCGGREDLGHLCMWSDCC RLRGRLQPMAGTCPAPERLABPLSCPPATTVA GCSRTPLDTVASSER GCSRDRSSTYLDLASHETMSPCACTRITT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT HONDLHVGSBLQVLVADTIDFCTELARCTINIT RCRGRSTRUTLYTTIMELTELBASQ GFPLETYRHFVTRIFFFRCSSLAGENVINSTYLCS LLLLSVVALFFYFTARCSLAGATSVICSTULLAB GGSBCRDLTALGGGREGATTTVAGERT RCRGRSTRUTLALGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTLAGGGREGATTTVAGERT RCRGRSTRUTL	1	Ì	1		, - •		
SIPPPGAGSYAYLALINSEARAGDKESQARSI, LARICLPHOAGGGBAKKLTBSYQLGLQCAGG AGRGDLYSRLVSVPPARERIFAVFERPQGSPA ARAPAALCAFRAPAVRAIRAARTAGYPEP APDVVAVLDSVVQGTGPACERKLINQLQPEP LDCGAAHQHPISLIQPKLATPVFRAPGLTSV AVASVNNYTAAVELPHINMQEPDABGYLY LMTSHQMARVKVAACNVHSTCGDCVGAAD AVCGWCALETRCTLQDCTNSSQHFWTNA SSGPSRCPAMTVLPSEDVRCPYFGMLQLGSS LPSISGMEMACDYOFONIRGYTARVPEPAGTISQ AYCGWCALETRCTLQDCTNSSQHFWTNA SGGPSRCPAMTVLPSEDVRCPYFGMLQLGSS LPSISGMEMACDYOFONIRGYTARVPEPAGTIQ IAYCNLLPRDQFPPFPPQDHVTVEMSVRVN GRNIVKANFTYPCSRTAQVYPTSNQD CPRTLLSPLAVTTGGSQNLLVPLANTAFFQG ALBCCSGQBEIGHCAWNSDVC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSQCLGREDIGHLCAWSDC RUNGSPDCSCREDIGHCAVCELP DRYTYSEEIVCVTGPAGPLSGVTVNASKE GSSBARSVPLPHALSEPTIGFRCAGTERIT HGNDLHVGSBLQVLVNDTDCTELMRTDTSI ACTIMEGCALPAPVCVGTFRAGCVTNASKE RUNGSQCSPLORIVENGTTLCCVUNSTLTCPSP GALSNASAPVDFTNGRYADEVAVAEELLD PEAQRGSSPRLDVLFVGTETTCCVVNFTLTCPSP GALSNASAPVDFTNGRYADEVAVAEELLD PEAQRGSSPRLDVLFVGTSTLTCSPTLTCVSCT RUNGSGSQCAGETHFLLGEVKSFRCWIKH HGCPLTLVHIVSTKGAGKEQDSLGLGSHEV RVKGGVSCDQIVSDGRIFICSVVESELGAAVGQ LPTTQVGNFNQTATLQLGGSETAITVSVTCSV RVKGGVSCDQIVSDGRIFICSVVESELGAAVGQ LPTTQVGNFNQTATLQLGGSETAITVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVSVTCSV VSCGRETIFICSVLFMCTTLVS			<u> </u>	<u> </u>	sequence		
LARICLPHÓAGGDAKKLTESYIQLGLQCAGG AGRGLYSRLVSVPARERITAVERPOGSPA ARRAPALCAFRADVRAAIRAARTACFVEP APDVVAVLDSVVQGTOPACERKLINIQLOPEQ LDCGAAHLQHPLSILQPLATTVFRAPGLTSV AVASVANTAVELGPTUNGRLIKENIQLESMQ VVSRRVYTVAYGEPVHHVMQFDPADSOTLY LIMTSHQMARVKVAACNYMSTCOPCVGAAD AYCQWCALETRCTLQQDCTNSSQQHFWTSA SEGFSRCPAMTVLPSEDVRQEFYGMILQISGS LPSLSGMEMACDYGRORIRTVARVPGAPGLTGSV GRNIVKAMFITYDCSSTAQVYPHTACTSCLSSA QWPCTWCSQQHSCVSNQSRCCASPPTTSPQD CPRILSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEEIFEAVWVNESVVRCDQVVLH TITKSQVPELSQLKGRARTLDSSPENTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLGGPLQPMAGTCPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELVCVTOPAPFRAIPEISQFLOGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLT WYNQNVSMAVHHIGRETLCXVLNSTILTCPS GRSRRSTVALPOTLANDELCSTLOGGT RCGRRCGRRCGRRCGGGGGGGGGGGT LLTIRGRAGGT RCGRRCGRRCGGGGGGGGGGGGGGGGGGGGGGGGGG	l	1	i		ł	ł	
AGROBLYSRL/SVYPARERL/BAYPERPOSPA RARAPALCARPRADVRAMIRACTOPE APDVVAVLDSVVQGTGPACERKI.NIQLOPE APDVVAVLDSVVQGTGPACERKI.NIQLOPE APDVVAVLDSVVQGTGPACERKI.NIQLOPE APDVVAVLDSVVQGTGPACERKI.NIQLOPE AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV AVASVNNYTAVPLGTVAGRILLKINILNENSIV GRANVKANFTTDCSTRAVVPTSVASVAN GRANVKANFTTDCSTRAVPTAGSQNILVPLANTAFFQ AVALESTOLEERFAVVNNESVYRCDQVVLIH TITKSQVPPLSLQLKGRRAFILDSPEMTVA VYNCAMGSPDCSGLOKEDLGFILCMWSDOC RIRGPLGPMAGTCPAPERRAEPILSGPLDGGT LLTRGRNIGRARISDVAGGVVTVNASSK GKSRDFSYVLPLVHSLEPTMGRAGGTRITT HGNDLHVGSELQVLVNDTDFCTELMRITDTSI ACTMEGGALPAPVVCVRERERGCVHGNLTF WYMQNPVITAISPRSPVSGGRTTTVAGERFH MVQNSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEAAGRGSRRIJVLYPHOPGSTAKREKWIKH HPGEPLIT.VHYSTIKGAGKEQDSLCLQSHEY RVKIGQVSCDIQIVSDRIHGENYAGERGVHGNLTF WYMGNSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEAAGRGSRRIJVSTORGEGVHGNLTF WYMGNSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEAAGRGSRRIJVSTORGEVHGNLTF WYMGNSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEAAGRGSRRIJVSTORGEVHGNLTH WYMGNSMAVHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEAAGRGSRRIJVSTORGEVHGNLTH HPGEPLIT.VHYSTIKGAGKEQDSLCLQSHEY RVKIGQVSCDIQIVSDRIHGSPTLGGRKTENSLAKEN WILL PEAAGRGSRRIJVSTORGEVHANT HTGREWENDLTTEVTHALGQOFFARME GISLFSSLLDNIKHTPFWCSSLGAWGG GIPELEVKHFVYTIFFFKCSSLYPSGOT LNSQGSSQAGETHPLI.GGRKTPSCSCPMGE GISLFSSLLDNIKHTPFWCASLGQNIKGSIDATUNIT KARYTINEEWLIJERNINSFOGGG MDSLSVRAMDTDTI.TTQVKEKLEAFCKNDH RCSLASLITIALHGKLEFYHJULLDDTYSV VEGGRKKLNTLAHYHTPHALEGQKFGASKGG MDSLSVRAMDTDTI.TTQVKEKLEAFCKNDH RCSLASLITIALHGKLEFYHJULLDDTYSV VEGGRKKLNTLAHYHTPHALEGQKFGASKGG MDSLSVRAMDTDTI.TTQVKEKLEAFCKNDH RCSLASLITIALHGKLEFTHALDLDDTYSV VEGGRKKLNTLAHYHTPHALGQGKFGGND RCSLASLITIALHGKARFTHALDLORG RCSLASHLITAHTATTGVARHER RCSLASHLITAHTATTGV	1	ļ					l
ARAPAALCAFREADVRAJERATRATACFVEP APDVVAUDSVQGTGACERUNIQQEQ APOVAVLDSVQGTGACERUNIQQEQ APOVAVLDSVQGTGACERUNIQQEQ APOVAVLDSVQGTGACERUNIQQEQ AVSONNYTAVGEPVHHIVMQCDPADSGYI, AVASVNNYTAVGEPVHHIVMQCDPADSGYI, LMTSHQMARVKVAACNVHSTGDCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEQPSRCPAMVILPSEDVRQEYPGMILQISGS LPSISGMEMACDYGNNIRTVARVCPAPCHQ IAYCALLPRQPFPFPNQDHVTVEMSWVNN GRNIVKANFITUNCSRTAQVYPHACTSCISA QWPCFWCSQQHSCVSNQSRCEASPNTTSDQ CPRTLISPLAPVPTGSSQNI UPLANTAFFQQ AALECSPGLEEFEAVWVNESVVRCDQVVLH TITKKSQVPFLSLQLKGRRARFLASTGLDGGT LLTRGRNLGRRI SDVAHGGWGGVACEPP DRYTVSEEIVCVTGAPCHASGYVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNNTDDFCHRRTDTSI ACTMPEGALPAPVVCVRFERRGCVHONLTY WYNGNPVTIAISPRSPYSGGRTHVAGERFH MVQNVSNAAVHHIGREPTLCKVLNSTLITCPSP GALSSAASAPVDFFNGRAYADEVAVAEELLD PEEAORGSRRLDVLPNPQPSTAKREKWIKH HPGEFLTLVHVSTKGAGKSGDGTAUSTVICOV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQUBEIERKGFAELQTIDMTDLTKELRISSQ GIPPLEVKHFVTTKGAGKSGDGTAUSTVICOV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQUBEIERKGFAELQTIDMTDLTKELRISSQ GIPPLEVKHFVTTFFFKCSSL YEERYVLPSQT LNSQGSSQAGETHPLLGEWKIPESCRPNME GISLFSSLDNKHFLIVEVHALEQQKDFAVRD RCSLASLITIALHGKLEYYTISMKELLVDLLD ASAAANPELMRRTESVVEKMLTNVMSICM YSCLRETVGEPFFLLLCAKQQNKGSIDATT KARYTINEEWLIKERNITNVSFQGCG MDSLSVRAMDTDTT.TQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTGSYTLADDTTSV VEGGRKKLNTLAHYKIPEGASLAMSLIDKKS, NTLGRVKDLDTEVFFHLLVPTLAELQCJCDSPSTN KLLYAKERPYFRLVQRYYKQIQDMTPLSQE MAHALAESSRYLVPFHLVTPTLBQDELAPFKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKEIDDLF KARYTINEEWLIKERNITNVSFQCGC MDSLSVRAMDTDTT.TQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTGSYTLADDTTSV VEGGRKKLNTLAHYKIPPGASLAMSLIDKKS NTLGRVKDLDTEXPFHLUPTBLAPFKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKEIDDLF KARYTINEEWLIKERNITNVSFGCG MDSLSVRAMDTDTT.TQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTGSYTLADDTTSV VEGGRKKLNTLAHYKIPPGASLAMSLIDKKAK RXTRPQMAALEAAPTARRTQLQHKFGDVVAL MEDNTYECYSEA MAHALAESSRYLOQAPDACSTSIDL.QGDSSTN KLLYAKERPSYRKUPRYNQLOCHDSTSPLT KLLYAKERPSYRKUPRYNQLOCHDATPLSQE MAHALAESSRYLOQREDATRONAETYYAK RXTRPQMAALEAAPTARRTQLQHKFGSVAC ARGEORGRNGSSGSPGELLAFERNLELQME ANGKONGRNGSGSPGFBLLAFERNLELQME THE TOWARDSTANDARD ANGKON		ŀ					
APDVVAVLDSVVQGTGPACERKLNIQLOPE LDCGAAHLQHPLSLQPKATPYFRAPGLTSV AVASVNNYTAVPLGTVAGRILKDNLNSBOWLY VARRVYTAVPLGTVAGRILKDNLNSBOWLY LMTSHOMARVKVAACKVHSTCGDCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGFSRCPAMTVLPSEDVRQFYFOMHLQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ IAYCGHLPROUPFPPPNQBHYTVSMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLS QWCCFWCSQQHSCVSNQSRC4SPPTTSQD CPRTLLSPLAPYPTGGSQNLVPLANTAFFQG AALECSPGLEEFFFAVWVNESVCHQVVLH TTRKSQVPPLSLQLKGRPARPLDSPEPMTVM VYNCAMGSPDCSCLGREDLGHCMVSDGC RLRGPLQPMAGTCPAFERAEPLSGPLDOGT LLTRRGNLGRRLSDVAGGVVGNGVACEPLP DRYTVSEBIVCVTGPAPGPLSGVVTVNASKE GKSRDFSVVLPLYHSLEPTMGPKAGGTEITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTIMPEGALPAPVPVCVRFERGCVCHGNLTF WYMQNPWTTAISPRSFVSGGRTTTVAGGRFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSSASAPVBFTNIKGAGKEQDSLCLGSHEY RVKIGQVSCDIQIVSDRIHGSVNESLGAAVQ LPITTQVGFNFOTIATLQLGGSETATUSVIVCSV LLLLSVVALFVFCTKSRRABRYWQKTLLQME EMESGNEERKGFAELQTDMTDTIXELINSSQ GPPLEVKHFVTRTFPFKCSSLYBERYVLPSQT LNSQGSSQAGETHTLGWENTSCLSVAGG LNSQGSSQAGETHLGEWKTPSCRPTMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVIVCSV LLLLSVVALFVFCTKSRRABRYWQKTLLQME EMESGNEERKGFAELQTDMTDTIXELINSSQ GPPLEVKHFVTRTFPFKCSSLYBERYVLPSQT RCSLASLTITALHGRESYTTSMERLLVDLID ASAAXNPKLMLRTESVVKKMLTNWASICM YSCLRETVGEPFFLLGEWKTPSCRPTMEE GISLFSSLDNKHFLIVFVHALEQQKDFAVING RCSLASLTITALHGRESYTTSMERLLVDLID ASAAXNPKLMLRTESVVKKMLTNWASICM YSCLRETVGEPFFLLGEWKTPSCRPTMEE GISLFSSLDNKHFLIVFVHALEQQKDFAVING RCSLASLTITALHGRESYTHJURTUSTCSLDTINS VEDGRKKLNTLATHYRDFGASLAMSLLDKKN NTLGRKVKDLDTEXYFFDTLEGEAPKKSH RQSHRKKVLPSTYTTRLISTKGTLQKFLDDTSV VEDGRKKLNTLATHYRDFGASLAMSLLDKKSH NTLGRKVKDLDTEXYFFTPLLCTRLDDTSV VEDGRKKLNTLATHYRDFGASLAMSLLDKNS PDTILHTWKTNSLPLRFWVNILKNPQFVFUDDT KLLYAKEPFYKRVQRYYKQIQDMTPLSSQE MAHALESSREYQHFMTNVAMAHYKYAK RYRPQMAALEAMPTARRTQLQHKFGQVSWPWQQ VANGKONQRNMGSPQPSLAFERRILELQM AENDTYFCYSEA RYRPQMAALEAMPTARRTQLQHKFGQVSWPWQQ VANGKONQRNMGSPQPSLAFERRILELDLAGL CAYSLLMGKLRPFVAKDTPTSTESHT LKD 730 2080 A 5747 1 382 PFLKCMRKAFRSSKLLQVGYTPDGKDDYRWC			1			[
LDCGAAHLQHPLSLQPKATPYFRAFGLTSY AVASVNNYTAVCHOTVGRULHNESMQ VVSRRVYTVAYCEPVHHIVMCPDPADSGYL LMTSHOMAR WKVAACNVISTGTGCVGAAD AYCGWCALETRCTLQDCTNSSQQHFWTSA SEGFSRCPAMTVLPSEDVRQEYPGMILQISSG LPSLSGMEMACDYGNNIRTVARAVPGAAFGHQ IAYCRLPROQFPFPPNODHVTVEMSVTW\ GRNIVKANFITUPCSSTAQVYPTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPNTTSGDC CPRTLLSPLAPPYTGSSQNLLYPLANTAFFQQ CAALECSGLEEIFEAVWVNESVVRCDQVVLH TITKKSQVPFLSLQLKGRPARRLSSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCAWNDGO CREGPLOPMAGTCPAPERLAEIPSGPLDGGT LLTRGRALGRRLSDVABGVVNGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVTTVASKE GKSRRFSYVLPLVHSLEPTMGFRAAGTRITI HGNDLHVGSELQVLVNDTDFCTELMRTDTSI ACTMEGGALPAPVCVRFRERGCVHGNLTF WYMQNPVTTAISPRSSPVSGGRTTTVAGERET MVQNSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFTNGRAYADEVAVAEELLU PERAGRGSFRLDYLPPNGFSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNSCLAGAAVQQ LPTTIQVGRFNOTTATLQLGGSETAIDVSTVCSU LLLJSVVALFVFCTKSRRAERYWQKTLLQME EMESQREEIRKGFAELQTIMTDLTKELINSQ GIPPLEYKIFVTRTFFFKCSSLYEERYVLPSQV LSQGSSQAQESTIPLLGGEWTEPSCRPNMEE GISLFSSLLDNKHPLLVFVHALEQQKDFAVNA RCSLASLLTTALHGKLEYYTSTMEELLVDLID ASAARNPRLMLRRTESVVEKMLTNWSFGQCG MDSLSVRAMDTDT.TQVKEKLECKONPY SQWPRAEDVDLEWFASSTGSYTLRDLDDTSV VEDGRKKLNTLARHVRPGGASLAGANSLDKSU NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK NTLGRVKDLDTEXYPHLVLPTDELAEPKSH RQSRRKKVLPEIYLTRLLSTKGTLQKFLDLSK RYRQDMAALEANFTARTQLQHKFGOVVAL MEDNIYECYSEA PTHLHWKTNSLPLRFWWILKNPQFVFDUA MAHALESSRKYONARHYSYNGQDMAHALESSRK RYRQDMAALEANFTARRTCQLJGKDGSTN KLLYAKEPYPKXVQRYYKQQDMTPLSGE MAHALESSRKYONARHYSTYSPLT LKD 730 2080 A 5747 1 382 FIKKCMKR						Į	
AVASVNNYTAVELGTVINGELKINI.NISMO VVSRRVYTAVELGTVINGELKINI.NISMO VVSRRVYTAVELGTVINGELKINI.NISMO VVSRRVYTAVELGTVINGELKINI.NISMO VVSRRVYTAVELGEVILQDCTNISGOFTWTSA SEGISRCPAMTVLPSEDVRQEVPGMILQISGS LPSLSGMEMACDYGNNIRTVARVPGAFGILQISGS LPSLSGMEMACDYGNNIRTVARVPGAFGILQISGS QWPCFWCSQQIBCS LPSLSGMEMACDYGNNIRTVARVPGAFGILQISGS QWPCFWCSQQIBCSCNOSRCSASPNPTSIQD CPRTILLSPLAPVPTIGGSQNILVPLANTAFFGG AALECSPGLEEFIEFAVWNNESVYRCDQVVLH TTRKSQVPPLSLQLIKGRRARFILDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHAWSDGC RLRGPLQPMAGTCPAPERIAEPISGPLOGG LLTIRGRNLGRISDVAIGGVVGCPELP DRYTYSEBLVCVTGPAPGPLSGVVTVNASKE GKSRDRFSVVLPLYHSLEPTMGPKAGGTEITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVREREPGVHONLTF WYMQNPVTAISPRSPVSGCRTTIVAGERIFH MVQRIVSMAVHHIGREPTLCKVLNSTLITCPSF GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRRLDVLPNGPSTAKREKWIKH HPGEFLILLVINGAGKKEDDSLGLGSHEY RVKIGQVSCDIQIVSDRIHIGSVNCSLGAAVGQ LPITIQVGFNFQTATALQLGGSETAVSVICSV LLLLSVVALFVFCTKSRRARRYWQKTLLQME EMESGNEERIKGFAELQTIDMTDTIKFLINRSQ GFPLEVKHFVTRTFFFKCSLYBERYVLPSQT LISQGSQAGSTHFLIGGWKTPSSCCFNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD GPPLEVKHFVTRTFFFKCSLYBERYVLPSQT LNSQGSQAGSTHFLIGGWKTPSCCFNMEE GISLFSSLDNKHFLIVFVHALEQQKDFAVRD RCSLASLITALHGKEPYTSIKMELVDLID ASAAXNPKLMLRTTESVEKMLTNMSISCUS LNSQGSQAGSTHLIGESVATSPERCH LOTTSV VEDGRKKLNTLAHVKIPSGASLAMSLIDKKN YSCLEFTVGFFFLILGEMKCPSCCFNMEE GISLFSSLDNKHFLIVFVHALEQQKDFAVRD RCSLASLITALHGKEPYTSIKMELDDTISV VEDGRKKLNTLAHVKIPSGASLAMSLIDKKN YSCLEFTVGFFTLLGAKKGDHAGDSTATG KARYTLNEEWILLRENIEAKPRNLNVSFQGCG MDSLSVRAMDITDTT,TQVEKELBAPCKNSY VEDGRKKLNTLAHVKIPSGASLAMSLIDKKN NILGKVKDLDTERYFHILVPTDELAPFKKSH RQSHRKKVIPSTYTRILISTKGTLQKFLDDTSV VEDGRKKLNTLAHVKIPSGASLAMSLIDKKN NILGKVKDLDTERYFHILVPTDELAPFKKSH RQSHRKKVIPSTYTRILISTKGTLQKFLDDTSV VEDGRKKLNTLAHVKIPSGASLAMSLIDKKN ARTSPORMATATARTQLQHKFEQVVAL MEDINTYCYSEA RYPQDMAALEANPTARTQLQHKFEQVVAL MEDINTYCYSEA RYPQDMAALEANPTARTQLQHKFEQVVAL MEDINTYCYSEA RYPQDMAALEANPTARTQLQHKFEQVVAL MEDINTYCYSEA RYPQDMAALEANPTARTQLQHKFEQVVAL MEDINTYCYSEA RYPQDMAALEANPTARTQLQHKFEQVVAL MEDINTYCYSEA PLILGKURDTYSTSPLT LKD	ŀ	,		Ì			APDVVAVLDSVVQGTGPACERKLNIQLQPEQ
VVSRAVVTVAYGEPYHIVMQFDPABSQTLY LMTSHQMARVEVAACNHSTGODCVGAAD AYCGWCALETRCTLQODCTNSSQQIFTYSA SEGFSRCTAMTVLSEDWQFYPGMLLQISGS LPSLSGMEMACDYGNNRTVARVPGPAFGHQ LAYCNLLPRDQFPFFPRQDHVTVEMSVRVN GRNIVKANTTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASFNPTSPQD CPRTLLSFLAPYFTGGSQNLY-LANTAFFQG AALECSFGLEEFFAVWVNESVVRCDQVVLH TTKKSQVFFLSLQLKGRPARLDSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAFERAFDSSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAFERAFDSSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAFERAFDSSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLGTMSDGC RLRGPLQPMAGTCPAFERAFDSSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLGTMSDGC RLRGPLQPMAHDLGSFETAVSSCACGEPL DRYTVSEEIVCVTGPAPGPLAFELSSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLGTMSDGC RLRGPLQPMAHDLGSFETAVSSCACGEPL DRYTVSEEIVCVTGPAPGPLASFLDSFEPMTVM VYNCAMSPLDSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLGTMSDGC RLRGPLQPMAHDLGSFETAVSSCACGEPL DRYTVSEEIVCATGCACGEPL DRYTVSEEIVCATGCACGEPL MVQNVSMAAVHIGREFTLCKVLNSTLITCRSSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDVIPMPGFRAGERTDVSTLTCTCVNSTLTCTCSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDVIPMPGGERGAVADEVAVAEELLD PEEAQRGSRFRLDVIPMPGGESCAGAVE RVKIGQVSCLQIQVSDRIHIECKVNSTLTCTCSVC LLLLSVVALFYCTCKSRRAGENSLGAAVQ QLPTTIQVGNNOQINATILOLGGSSCTAUVSTVCSV LLLLSVVALFYCTCKSRRAGENSLGAAVQ QLPTTIQVGNNOQINATILOLGGSSCTAUVSTVCSV LLLLSVVALFYCTCKSRRAGENSUSLGAAVQ QLPTTIQVGNNOQINATILOLGGSSCTAUVSTVCSV LLLLSVVALFYCTCKSRRAGENSUSLGAAVQ QLPTTIQVGNNOQINATILOLGGSSCTAUVSTVCSV LLLLSVVALFYCTCKSRRAGENSUSLGAAVQ QRFTRLTTVHTVSTKGGCSTATVSTVCSV LLLLSVVALFYCTCKSRRAGENSUSLGAAVQ QRFTRLTTVHTVSTKGGCSTATVSTVCSV LLLLSVVALFYCTCKSRRAGENSUSLGAAVQ QRFTRLTTVHTVTTRTFFFCCSLTGATAT RVGGVSCQAGGTHTLLGGWTGSTCAG RSCRAGENTAMT RVGGVSCQAGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT]	Ì				}	LDCGAAHLQHPLSILQPLKATPVFRAPGLTSV
LIMTSHQMARVKVAACNVHSTICGDCVGAAD AYCGWCALETRCTLQQDCTINSSQQHFWTSA SEGFSRCPAMTVLPSEDVRQFYPGMILIGIGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ IAYCNLLPRQGPPFPPRODHVTVEMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCASFNPTSSQD CPRTLLSPLAPVPTGGSQNILVFLANTAFFQG AALECSFGLEEFFAVWVNSCEASFNPTSSQD CPRTLLSPLAPVPTGGSQNILVFLANTAFFQG AALECSFGLEEFFAVWVNSCEASFNPTSSQD CPRTLLSPLAPVPTGGSQNILVFLANTAFFQG AALECSFGLEEFFAVWVNSCCAPLP DRYTVSEBIVCVTQPAFGBLGHLCMWSDGC RLKGPLQPMAGTCPAPERAIEPLSGPLDGGT LLTTRGRNLGRRLSDVAHGVNIGQCFLP DRYTVSEBIVCVTQPAFGBLGHLCMWSDGC RLKGPLQPMAGTCPAPERAIEPLSGPLDGGT LLTTRGRNLGRRLSDVAHGVNIGNLTF WYMQNPVTLSISPTFMGFKAGFRTITH HGNDLHVGSELQVLVNDTDCTELMRTDTSSI ACTMPEGALPAPVPVCVKFERRGCVFIGNLTF WYMQNPVTLSIPPRFSSGRTITVVAGERFH MYQNVSMAVHHIGREFTLCKVLNSTLITCPSP GALSNSASPVDFFTNGRAYGPVAFELLD PEEAQRGSFFLDVLPPPOFSTAREEKWIKH HPGBPLTLVHVSTKGAGKEQDSLCLQSHEY RVKIGQVSCDIQIVSDTHIECSVPASELGAVGG LPTTQVGNNNQTATLOLGGSSTAILVSNICSV LLLLSVVALFYFCTKSRRAERXWQKTLLQME EMBSQREERKGFAELQTDMTDLTKSLNSQ GIPPLEVKHTVFTKFRKGSLVERSVVLSDLD ASAAKNPKLMLRRTTSVVEKMLTNWMSICM YSCLEETVGEFFELLLCAIKQDINGGSDATIO KARYTLNEEWLLRENIEAKRRNLNNYSPGGC MDSLSVRAMDTDLTQVKEKLEAFCKNNYY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKINTLAHYKIPEGALASLIJDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHFRKVLPFTLYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RQSHRKKVLPFTYTLLISTALASLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAPPKKSH RXHPQMMALEANPTARTQLQHKFGSVFVAK RYPQMAALEANPTARTQLQHKFGSVFVAK RYPQMAALEANPTARRTQLQHKFGSVFVAK RYPQMAALEANPTARRTQLQHKFFTYAK RYPQMAALEANPTARRTQLQHKFFTYAK RYPPQMAALEANPTARRTQLQHKFFTYAK RYPPQMAALEANPTARRTQLQHKFTYAK RYPPQMAALEANPTARRTQLQHCHFTYAK RYPPQMAALEANPTARRTQL	ĺ						
AYGWALETRCTLQOCTNISSQOHYNTAS SEGPRCAMTVLPSEDVROSPYCHMLQISCS LPSLSGMEMACDYGNIRTVARVPGPAFGHQ IAYCHLLPRDGFPPPNQDHYTVEMSVRVN GRNYKANTITYDCSRTAGVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASFNTTSQD CPRTLISPLAPVFTGGSQNILVANTAFFQG AALECSFGLEEIFEAWVNESVVRCDQVVLH TTRKSQVPPLSLQLKGRPARPLDSFEPMTVM VYNCAMGSPDCSQCLGREDLGHLCAWSDCC RLRGQPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLITEGRNLGGRLSDVAHGWVIGGVACEPLP DRYTVSEEIVCVTOPAPGPLSGVTVTNASSE GKSRDRFSVLPLHVISLEFTHMGFKAGGTRITT HGNDLHVGSELQVLVNDTDPCTELMRITDTSI ACTMPEGALPAPVPVCVEFERGCVHGNLTT WYMQNPVITAISPRSSPVSGGRITTVAGEREH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVALDE PEEAQRGSRFRLDVLPNPQFSTAKREKWIKH HPGEPLTLVHIVSTKGAGKGDSLGLGSHEY RVKIGGVSCDIQIVSDRIHICSVNESIGAAVGQ LPITTIQVGNNTQNITATLQLGGSTAATVSIVICSV LLLLSVVALFVFCTKSRRAERTWQKTLQME EMESQREEIRKGFABLQTDMTDLTKLLDNSQ GIPFLEYKHFVTRIFFFKCSSLYEERYVLPSQT LNSQGSSQAQETHFLLGEWKIFESCRYNMEE GISLFSSLLDNKHFLJVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEHLAEQQKDFAVRD RCSLASLLTIALHGKLEYTSISMKELLVDLID ASAAKNPKLMLRRTESVVEHLAEQGKDFAVRD RCSLASLLTIALHGKLEYTSISMKELLVDLID ASAAKNPKLMLRRTESVVEHLAEQGKDFAVRD RCSLASLLTIALHGKLEYTSISMKELLVDLID ASAAKNPKLMLRRTESVVEHLAEQGKDFAVRD RCSLASLLTIALHGKLEYTSISMKELLVDLID ASAAKNPKLMLRRTESVVEHLAPCKNVPY SQWPRAEDVDLEWFASSTQSYLIROLDDTSV VEDGRKKLNTLAHQVEHCASSTQSYLIROLDDTSV VEDGRKKLNTLAHYLDEGALMSLDKKD NTLGRVKDLDTTEKYTPHLVLPTDELAEPKSH RQSHRKKVLPFTYLTILLISTRATJVSPQGCG MDSLSVRAMDTDTLTQVKEKILEAPEKKSH RQSHRKKVLPFTYLTILLISTRATJVSPQGCG MDSLSVRAMDTDTLTQVKEKILEAPEKKSH RQSHRKKVLPFTYLTILLISTRATJVSPQGCG MDSLSVRAMDTDTLTQVKEKILEAPEKKSH RQSHRKKVLPFTYLTILLISTRATJVSPQGCG MDSLSVRAMDTDTLTQVKEKILEAPEKKSH RQSHRKKVLPFTYLTYLLTSTRATJQCHKKFS VEDGRKKINTLSPTYTJRLTREPTYLNSPQFVDDIX RKLLYAREPYRKVQVQTYKQQDMTJSEQE MNAHLAEESSRKYQNEFTNIVAMAEETKYAK RYRPQMAALEANTTARRTQLQHKFFQVVVAL MEDNIYECYSEA POPSPLFHSHLETTQLLRTAQLPEQVSWPWQQ VANGKGNGRNAMGSPQPSLLAFERNLELQMG LGYSLLMGKLRPRVAKDTLRVHRDSTTSPLT LKD 2880 A 5744 3 292 VANGKGNGRNAMGSSPQPSLLAFERNLELQMG LGYSLLMGKLRPRVAKDTLRVHRDSTTSPLT LKD VANGKGNGRAMGSPQPSLLAFERNLELQMG							VVSRRVVTVAYGEPVHHVMQFDPADSGYLY
SEGPSRCPAMTV12SEIDVRQEYPGMIQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ LAYCNLLPRDQPPFPPPNQDHVTVEMSVRVN GRNIVKANFITYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCASFNPTISNQD CPRILLSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEEFEAVWVNEXCDQVVLH TTRKSQVFFLSLQLKGRPARFLDSFFFMTVM VYNCAMGSPDCSQCLGREDIGHLCMWSDGC RLRGPLQPMAGTCPAFERAIEPLSGPLDGGT LLTEIGRNLGRR.SDVAHGWGVACCELP DRYTVSEEIVCVTGPAFGPLSGVLTOGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPBGALPAPVPVCVRFERRGCVHGNLTF WYMQNSVTAISPRRSPVSGGRITITVAGERFH MVQNVSMAVHHIGREFTLCKVLNSTLITCPSP GALSNASAPVDFFTNGRAYAVAEELLD PEEAQRGSRFRLDVLPPPQFSTAKREKWIKH HPGGPLTLVHYSTKGAGEGSGTITVAGERFH MVQNVSMAVHHIGREFTLCKVLNSTLITCPSP GALSNASAPVDFFTNGRAYAVAEELLD PEEAQRGSRFRLDVLPPPQFSTAKREKWIKH HPGGPLTLVHYSTKGAGEGSGLIGJSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVQQ LPTTQGNPNGTATLQLGGSSTAITVSTICSV LLLLSVVALFVFCTKSBRAERFWQKTLLQME EMBSQREERKGFAELQTDMTDLTKSLNSQ GIPFLEVKHYVTSTFFRCSSLSLEGSVLVALPSQL GIPFLEVKHYVTSTFFRCSSLSVERYVLPSQT LNSQSSSQAQETHFLLGEWKIPSSCRTNMEE GISLFSSLLDNKHFLVFVTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKALBQCDFAVRD RCSLASLLTTALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKALBQCDFAVRD RCSLASLLTTALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKALBAGCNDFAVRD GRIPFSTLTLAHGKLEYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKALBAGCNDFAVRD RCSLASLLTTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKALBAGCNDFAVRD RCSLASLLTTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKALBAGCNDFAVRD RCSLASLLTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKALBAGCNDFAVRD RCSLASLLTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKALBAGCNDFAVRD RCSLASLLTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKLLBAGCCDFAVRD RCSLASLLTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKLLBAGCCDFAVRD RCSLASLLTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKLLBAGCCDFAVRD RCSLASLLTALHGKLEYTSISMEELLVDLID ASAAKNPKLMLRRTESVVEKLLBAGCCDFAVRD RCSLASLLTALHGKLEYTSISMEELLVDLID ASAAKNPKLMRTSSVEKLLQVGYTPDEAGERCNDPT RCSLASLLTALHGKLEYTSISMEELLCQMG PTHTHIVKTNSLTPRFTURLGCRTNAME RCSLASLTTALHGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC			ļ				LMTSHQMARVKVAACNVHSTCGDCVGAAD
LPSLSGMEMACDYGNINRTVARYPGAPCHIQ LAYCNLIPROPPEPPRONDHVTVEMSVRYN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA QWPCTWCSQQHSCVSNQSRCASSPITSISQD CPRTLISPLAPVPTIGSSQNILVPLANTAFFQG AALECSFGLEEFEAVWVNESVYRCQVVLH TTRRSQVPFLSLQLKGRFARFLDSPEPMTVM VYNCAMGSPDCSQCLGRBDLGHLGMWSDGC RLRGPLQPMAGTCPAPFREPLSQPLDGGT LLTIRGRNLGRRLSDVAHGVWIGQVACEPLP DRYTVSEEIVCVTGFAPCPLSGVVTVNASKE GKSRDRRSVYLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDDCTELMRTDTSI ACTMFEGALPAPPVCVYFERRGCVHGNLTI WYMQNPVTTAISPRSPYSGGRTTITI USGSSQAQBTIPILITI USGSSQAQBTIPILITI WYMQNPVTTAISPRSPYSGGRTTITI USGSSQAQBTIPILITI USGSSQAQBTIPILITI USGSSQAQBTIPILITI USGSSQAQBTTIPILITI US	l .	}	}				AYCGWCALETRCTLQQDCTNSSQQHFWTSA
I AYCNLLPRDQPPPPPNQDHTVTEMSVRYN, GRNIVKANFITYDCSRTAQVYPHTACTSCLSA QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEFEFAVWNESVVRCDQVVLH TTRKSQVFPLSLQLKGRPARFLDSPPMTVM VYNCAMGSPDCSQCLGRBDLGHLCMWSDGC RLRGPLQPMAGTCPAPERAEPLSQPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSELIVCVTGPAPQPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDCTELMRTDTSI ACTMPEGALPAPVPVCVRFERGCVHGNLTF WYMQNPVTLAISPRRSPVAVAEELLD PEEAQRGSRRLDYLPNOFSTAKREEWIGH HPGEPLTLVHVSTKGACKEQDSLGJSHEY RVKIGQVSCDIQVSDRIHGSVVESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAITVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTILQME EMESQRFEBRKGFAELQTIDMTDLTKELNRSQ GIPFLEYKHFVTRTFFFKCSSLYERRYVLPSQT LNSQGSSQAQETHPLLGAKHPESCRYNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTALHGKLEVFISMKELLVDLD ASAAKNPKLMLRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLCAKQQNNKSDDATI KARYTLNEEWULERNEAKFRNLNYSPGGC MDSLSVRAMDTDTLTQVKEKLLARCKNNPY SQWPRAEDVDLEWFASTGSYTLRDLDDTSV VEDGRKKLNTLAHVKPBGASLAMSLDKKD NTLGRVKDLDTEXYFHLVLPTDELAEPKSH RQSHRKKVLPETYLTRLISKGTLQKFLDDLF KALLSREDKPPLAVKYFPDFLEQAEKGISD PDTLHWKNTSLPLFFWNILKNPGVFPDIDK TDHDACLSVIAQAFDACSISDLQLGKDSTTN KLLYAKEBETYRKIVQRYGODMTPLSEQE MNAHLAEESRKYQNEPNTNVAMAEIYKYAK RYRPQMAALEANPTARRTQLQHKFEQVVAL MEDNIYESCYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLATAQLPEXYSPWGO TANGKGNGRNGSSPGPSLLAFERNLELQMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD TJLDLAKTRONGRAMSPORT			1	1			SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS
GRNIVKANFTIYDCSCRAGNTAGVYPHTACTSCLSA QWPCFWCSQQHSCVSQQSCCASPNTSPQD CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG AALECSPGLEEFEAVWYNESVYRCQVVLH TTRKSQVPLSLQLLGRFARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAPERLEPLSQPLDGGT LLTRGRNIGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGTSQVTVTNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELRTDTISI ACTMPEGALPAPPVCVFRERGCVHGNLTF WYMQNSWMAVHHIGREFTLSQVLSGNILTTCPSP GALSNASAPVDFFINGRAYADEVAVAELLD PEEAQRGSRFLDYLPNOYSTARREWWICH HPGEPLTLVIHVSTKGAGKEQDSLGSHEY RVKIGQVSCDIQIVSDRIHCSVNESSLGAVGGQ LPITIQVONFNQTIATLGGSETAITVSIVICSV LLLLSVVALFVFCTKSRRAERYWGKTLLQME EMESQURESIRKGFAELQTDMTDLTKELNSQ GIPFLEYKHFVTRTFFFKCSSLYERRYVLPQOT LNSQGSSQAGETHPLLGGSETAITVSIVICSV LLLLSVVALFVFCTKSRRAERYWGKTLLQME EMESQURESIRKGFAELQTDMTDLTKELNSQ GIPFLEYKHFVTRTFFFKCSSLYERRYVLPQOT LNSQGSSQAGETHPLLGAREPSKLTDAWD RCSLASLLTIALHGKLEYYTSIMKELVDLID ASAAKNFKLMRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQNKSGDATI KARYTLNEEWULRBRMESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQONKSGDATI KARYTLNEEWULRBRMESVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQONKSGDATI KARYTLNEEWULRBRMESVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQONKSGDATI KARYTLNEEWULRBRMESVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQONKSGDATI KARYTLNEEWULRBRMESVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQONKSGDATIO KARYTLNEEWULRBRMESVERMLTNWMSICM YSCLRETVGEPFFLLLCAIKQONKSGDATIO KARYTLNEEWULRBRMESVERMLTNWMSICM YSCLRETVGEPFFLLLCAIKQONKSGDATIO KARYTLNEEWULRBRMESVERMLTNWMSICM YSCLRETVGEPFFLLCAIKQONKSGDATIO KARYTLNEEWULRBRMESTERVLYNETYTHLEAPKKSH RQSHRKKLYLFENTLAIKLIKNPGVFPOIDK THIDACLSVIAQARDACSISDLOLGKDSFTN KLLYAKEIPEVREVVRYYKYQOLMFTLSEQ MNAELABESSKLYQNEFTNNVAMAENKYAK RYRPQIMAALEANPTARRTQLQHKKPGVVAL MEDDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLETAQLPEGVSPWGG VANGKGNGSOPGSPOSPLAFERNIELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD TILLD TIL							LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ
QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD CPRILSPLAPVPTGGSQNILVPLANTAFPQG AALECSFGLEEIFEAVWNESVVRCDQVVLH TITRKSQVFPLSLUKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAPEIRAIEPSGPLDGGT LLTIRGRNLGRRLSBVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVUTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMKTDTISI ACTIMEGGALPAPVPVCVRFERRGCVHGNLTF WYMQNVSNAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDPFINGSPAGGGRTTIVAGERFH MYQNVSNAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGSVAGERFH MYQNVSNAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGSVAGERFH WYGNVSDAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGSVAGERFH WYMQNVSNAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGSVAGERFH MYQNVSNAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFTRIFISCVNESSLGAAVGQ LPITTQVGGNFNQTIATLQLGGSETAITVSTVICSV LLLLSVVALFYFCTKSARAERYWQKTLLQME EMESQREEIRKGFABLQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFFKCSSLYEERTVLFSVILQME EMESQREEIRKGFABLQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFFKCSSLYEERTVLFSVILQME EMESQREEIRKGFABLQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFFKCSSLYEERQLGDATG RCSLASLITALHGRUFTYSTIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEFFFILLCAIKQQNKGSIDAITG KARYTILNEEWLLREIBLAKPRNLNVSPQGCG MDSLSVRAMDTDTLTQVKEKLLEAFCKNVPY SQWPRAEDVDLEWFASTQSVILRDLDDTSV VEDGRKKLNTLAHYKPPGASLAMSLIDKKD NTLGRVKOLDTEKYFTVSIMKELLDDTSV VEDGRKKLNTLAHYKPPGASLAMSLIDKKD NTLGRVKOLDTEKYFTVSIMKELLDATEN RCSLASLITALHGRUFTVFNOLLKNPQFVFDDK TDHIDACLSVIAQAFIDACSISDLQLGKDSFTN KLLYAKEPFYKRIVQTYKQIQDMTPLSQE MNAHLAESSRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNTYECYSEA MAHLAESSRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNTYECYSEA MAHLAESSRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNTYECYSEA MAHLAESSRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNTYECYSEA MAHLAESSRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNTYECYSEA MAGNGRNMAGSPSLLAFERNLELQDMG VANGKGRORMAGSPSLLAFERNLELQDMG VANGKGRORMAGSPSLLAFERNLELQDMG VANGKGRORMAGSPSLLAFERNLELQDMG VANGKGRORMAGSPSLLAFERNLELQDMG VANGKGRORMAGSPSLAFFSPSLLAFERNLELQDMG VANGKGRORMAGSPSLAFFSPSLLT TOTAL							
CPRTILSPLAPVPTGGSQÑILVPLANTAFFQĞ AALECSFGLEEIFEAVWVNESVVRCDQVVLH TITRKSQVFPLSLQLKGRPARPLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGG RLGGLQPMAGTCPAEIAIPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRIDTISI ACTIMEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MYQNVSAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDVLPPNPCSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLGSHEY RVKIGQVSCDIQVSDRIIHCSVNESLGAAVGQ LPHTIQVGFNOQTIALOGGSETAHVSVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIDMTDLTKELNRSQ GIPPLEYKHFVTRIFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLDNKHEFLYVALLEQQKDGSDATO KARYTLNEEWLLACHKOONNGGSDATO KARYTLNEEWLLACHKOONNGGSDATO KARYTLNEEWLLACHKOONNGGSDATO KARYTLNEEWLLACHKOONNGGSDATO KARYTLNEEWLLACHKOONNGGSDATO KARYTLNEEWLLACHKOONNGGSDATO KARYTLNEEWLLACHKOONNGGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSDATO KARYTLNEEWLLACHKOONNGSPTILACHKOONNGSPTILORPCKSLE GSSIRKKVLPETYLTRLLSTKGTLOKFLODDTSV VEDGRKKLNTLAHYKPPGASAMSLLOKKO NTLGRVKOONNGTHAGSPTILACHKOONNGTHAGSPCVSUPWOO MOSLSVAKANDTDTLTQKKELLEAPCKNOPY SQWPRAEDVOLLEWFASSTQSYILROLDDTSV VEDGRKKLNTLAHYKPPGASTQVIKWOONNGTHAGSPCVSUPWOO TOHDDACLSVLAQAFIDACSISDLOGKDSPTIN KLLYAKEPSPKKIVORYTYRQIODMTPLSQE MNAHLAEESRKYQNEPNTNVAMAENTCYRA RYRPQIMAALEANPTARTQLQHKFEQVSWPWGO VANGKGNORMMSGSPSPSLLAFERNLELQDMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD T30 2080 A 5744 3 292 QFSFLFHSHLETTQLLTRTAQLPEQVSWPWGO VANGKGNORMMSGSPSPSLLAFERNLELQDMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD T31 2081 A 5747 1 382 FLKCMKKAFRSSKLLQVGYTPDGKDDYNLON	j						GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA
AALECSFGLEEIPEAWWNESVYRCDQVVLH TTRKSQVPPLSIQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAPERALPLSGPLDGGT LLTIRGRNIGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMRPKAGGTRITI HONDLHVGSELQVLVNDTDPCTELMRIDTISI ACTMPEGALPAPVPVCVRFERGCVHGNLTF WYNQNPVTNISPRSFVSGGRITTVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFTNGRAYADEVAVAEELLD PEEAQRGSRFRLDVTNGPSTAKREKWIKH HPGBFLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQVVSDHIPCSVAESLGAAVGQ LPITIQVGNFNQTITATLQLGGSETAHVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQREERKGFAGLOTDMTDLTKELWSQ GIPPLEYKHPVTRTFFPKCSSLYEERYVLPSQT LNSQGSQAQETHPLLGGWEIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQDKGSIDAITG KARYTINEEWLRENIBARRNINVSPQGCG MDSLSVRAMDTDTLYCKEKLLEAFCKNVPY SQWPRAEDVDLEWFASSTGSYTRDLDDTSV VEDGRKKLNTLAHYBGSALAMBLDKLD NTLGRVKDLDTEKYPHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KARYTINEEBCKPHLAVKYFFDFLEQAEKGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEPFYRKVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMALEANPTARTQLQHKFEQVSWPWGQ VANGKGNQRMMGSPGSPLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5744 3 292 QPSFLFHSHLETIQLLRTAQLPEQVSWPWGQ VANGKGNQRMMGSPGPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD			İ]			QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD
TTRKSQVPPLSLQLKGRPARFLDSPEPMTVM VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTRGRNLGRRLSDVABIGVWIGGVACEPLF DRYTVSEEIVCVTIGPAGPLSGVTVTMASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRITDTSI ACTMPEGALPAPVPVVRFRERGCVHGNLTF WYMQNPVITAISPRRSDVSGGTITTVAGERFI MYQNYSMAVHHIGGRETLICKVLNSTLTCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPGPSTAKREKWIKH HFGEFLIT.VHVSTKAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ LPTITQVGNFNQTIATLOGSGSTAITVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDDTDTDLTKELNRSQ GIPPLEYKHFVTRIFPFKCSSLYEERYVLPSQT LNSQGSSQAQETHFLLGEWKIPSSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNMEE GISLFSSLLDNKHFLILGEWKIPSCRPNWEICM NSLSVARADTIDTLICGESCRATIVIVSCRP RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKKILTNMSSICM YSCLRETVGEFFTLLLGEWKIPSCRPNWEICM KARYTINEEWLLRENIEAKFRNINVSPQGCG MISLSVARADTIDTLICAKQQDINGSIDAITG KARYTINEEWLLRENIEAKFRNINVSPQGCG MISLSVARADTIDTLICAKQQDINGSIDAITG KARYTINEEWLLRENIEAKFRNINVSPQGCG MISLSVARADTIDTLICAKQDINGSIDAITG KARYTINEEWLLRENIEAKFRNINVSPQGCG MISLSVARADTIDTLICAKQDINGSIDAITG KARYTINEEWLLRENIEAKFRNINVSPQGCG MISLSVARADTIDTLICAKQDINGSIDAITG KARYTINEEWLLRENIEAKFRNINVSPQGCG MISLSVARADTIDTLICAKQDINGSIDAITG KARYTINEEWLLRENIEAKFRNINVSPQGCG MISLSVARADTIDTLICAKQDINGSIDAITG KARYTINEEWLLAGERGISD PDTLHIWKTNSLPRFWNILKNPSTQGCG MISLSVARADTIDTLICAKQDINGSIDAITG KARYTINEEWLLAGERGISD PDTLHIWKTNSLPRFWNILKNPSTQGCG MISLSVARADTIDTLICAKQDINGSIDAITG KARYTINEEWLAGERGISD PDTLHIWKTNSLPRFWNILKNPSTQGCG MISLSVARADTIDTLICAKQDINGGTAGGSPGTLAFGRNICGSPGG WANGKGROMMGSPGSPGLAFGENCHOMG VANGKGROMMGSPGSPGSLAFGENCHOMG VANGKGRO		1		1			CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG
VYNCAMGSPDCSQCLGREDLGHLCMWSDGC RLRGPLQPMAGTQPAERAISPLSQPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVYTVNASKE GKSRDRFSVYLPLVSLEPTMGFKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRSPYSGGRITIVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFRRAVAFELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSGNIHCSVNESLGAAVGQ LPTTIQVGNPNQTTATLQLGGSETAITVSTVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQREEIRKGFAELQTDMTDLTKELMRSQ GPPTLEYKHFVTRTFFFRCSSLYBERYVLPSQT LNSQGSQAQETBPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTISMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQNKGSIDAITG KARYTILNEEWLLRENIEAKPRNLNVSPQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTGSYLRDLDDTSV YEDGRKKLNTLAHVKSPGGSTANSLDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEYLTRLLSTKGTLQKFLDDLF KALSIREDKPPLAVKYFTPDLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQPVFDDIK TDHDACLSVIAQAFAXYFTPDLECQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQPVFDDIK TDHDACLSVIAQAFTARTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNNGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	1		i				AALECSFGLEEIFEAVWVNESVVRCDQVVLH
RLRGPLOPMAGTCPAPERAEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTIMGPKAGGTRITI HGNDLHVGSEL_QVLVNDTDPCTELMRIDTISI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PPEAQRGSRFRLDYLPNPQPSTAKREKWIKH HPGEPLTLVIHVSTKGAGKEQDSLGQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQREEIRKGFAELQTDMTDLTKELNRSQ GPPLEYXHFVTRIFFYCSSL YEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQNIKGBDAITG KARYTINEEWLLRENIEAKFRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYLRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEXYFHLVLPTDELAEPKKSH ROSHRKVLPFEYITRLLSTKGTLOKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVINLKNPQPVFDDIK TDHIDACLSVIAQAFIDACISIDLQLGKDSPTN KLLYAKEPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFINIVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVSUL MEDNIYECVSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNNGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	1		l				TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM
LLTIRGRNILGRRI SDVAHGWWIGGVACEPLP DRYTVSEIVCVTGPAPGPLSGVTVNASKE GKSRDRFSYVLPLVHSLEPTMGFKAGGTRITI HGNDLHVGSELQVLVNDTDCTELMRITDTSI ACTMPEGALPAPVPCVRFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCESP GALSNASAPVDFFINGRAVADEVAVAEELLD PEEAQRGSRFRLDYLPPPOPSTAKREKWIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGQ LPITIQVGNFNQTTATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTSSAYEARAFY WQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPPLEYKHFVTRIFFFKCSSLYEERYVLPSQT LNSQGSQAQETHPLLGEWKIPESCRPINMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQNKGSIDAITG KARYTLMEEWLIERMEAKPRNLNVSPOGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTGSYLRDLDDTSV YEDGRKKINTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEXYPHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KALSIREDKPPLAVKYFFPPLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQPVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSFTM KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESSKYQNEFINIVAMAEIYKYAK RXFPQIMAALEANTARTQLQHKFEQVVAL MEDNIYECVSEA 730 2080 A 5744 3 292 QPSPLFHSHEETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	ĺ						
DRYTYSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRIDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNNSMAVHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFTNGRAYADEVLAD PEEAQRGSRFLDVLPNPQFSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGQ LPITTQVGNFNOTIATLOLGGSETAITVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTIMOTIDLTKELNRSQ GIPFLEYKHFVTTRIFFFRCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLM.RRTESVVEKMI.TNVMSICM YSCLRETVGEFFFLLLCAIKQQINKGSIDAITG KARYTLNEEWILRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYLIRDI.DDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLUPTDELAEPKKSH RQSHRKKVLPETYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PPITLHHWKTNSLPLRFWNILKNPQFVFDIDK TDHIDACLSVLAQAFIDACSISDLOLGKDSPTN KLLYAKEPEYRKIVQRYYKQQDMTPLSEQE MNAHLAEESRKYQNEPNITNVAMAEIYKYAK RYXPQIMAALEANPTARRTQLQHKFPEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGRORRMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	l]				RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT
GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDDCTELRRIDISI ACTMPEGALPAPVPVCVRFERGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREETRKGFAELQTIDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDINKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFTLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENEAKPRNLNVSPQGCG MDSLSVRAMDTILTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVEDLDTEKYPHLVLIPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KALSREDKPPLAVKYFTDFLEQAEKRGISD PDTLHWKTNSLRFWVMILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEPEYRKIVQRYYKQQQDMTPLSQCE MNAFILAEESRKYOQRYYKQQQDMTPLSQCE MNAFILAEESRKYOQRYYKQQQDMTPLSQCE MNAFILAEESRKYOQRYYKQQQDMTPLSQCE MNAFILAEESRKYOQRYYKQQQDMTPLSQCE MNAFILAEESRKYOQRYYKQQQDMTPLSQCE MNAFILAEESRKYOQRYYKQQQDMTPLSQCE MNAFILAEESRKYOQRYMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD	1	ĺ	ŀ				LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP
HGNDLHYGSELQVI_VNDTIPPCTELMRITDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLIT WYMQNPVPUTAISPRRSPVSGGRTITVAGERFH MYQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSFRLDYLPNPQPSTAKREKWIKH HPGEPLTLVIHVSTK.GAGKEQDSL.GLQSHEY RVKJGQVSCDIQIVSDRIHCSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTKSBRAERYWQKTLLQME EMESQIREERIKGFAELQIDMTDLTKELLNRSQ GIPFLEYKHFVTRIFFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPMMEE GISLFSSLLDNKHFLLVFVHALEQQKDFAVRD RCSLASILTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEFFLILCAKKQQNKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEEKLEAFCKNVPY SQWPRAEDVDLEWFASSTGSYMLROLDDTSV VEDGRKKLNTLAHYKIPGGASLAMSLIDKKD NITLGRVKDLDTEKYFFHLVLPTGLAEPKKSH RQSHRKKVLPETYLTRLLSTKGFLQKFLDDLF KAALSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVLQAAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				ļ			DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE
ACTMPEGALPAPVPVCVRERRGCVHGNLTF WYMQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEGAQRGSRFRLDVLPNPQPSTAKREKWIKH HPGEPLTLVHVSTTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHICSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAITVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREERKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRIFFFKCSSL YEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASSLITIALHGKLEYYTSIMKELLVVLDID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLREINEAKPRINVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NITLGRVKDLDTEXYFHLVLIPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAALSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHTWKTNPTHLVRPFQFODK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEGSKQNEFINTVAMAEITYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNILELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC			ŀ	!			GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI
WYMQNPVITAISPRESPVSGGRITIVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITICFSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIHICSVVESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQIREIRKGFAELQIDMTDLTKELNRSQ GIPFLEYKHFVTRIFFPKCSSLYEERYVLFSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTINEEWLLRENIEAKPRNINVSFQGCG MDSLSVRAMDIDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKPPGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPLAVKYFFPDFLEQAGKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVPDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRLVQNYFVQIQDMTPLSEQE MNAHLAEESSRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QFSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							HGNDLHVGSELQVLVNDTDPCTELMRTDTSI
MVQNVSMAVHHIGREPTI.CKVI.NSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYL.PNPQPSTAKREKWIKH HPGEPLTLVIHVSTK.GAGKEQDSL.GLQSHEY RVKIGQVSCD.IQIVSDRIHIGSVNESI.GAAVGQ LPITIQVGNFNOTIATLQL.GGSETAINUSIVICSV LLLLSVVALFVFCTK.SRRAERYWQKTLLQME EMESQREEIRK.GFAELQTIDMTDLTKEL.NRSQ GIPFLEYK.HFVTRTFFPK.CSSL.YEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISL.FSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTALHGKLEYYTSIMK.ELLVDLID ASAAKNPKLMLRRTESVVEK.ML.TNWMSICM YSCL.RETVGEPFFLLLCAIKQQINKGSIDAITG KARYTI.NEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDITDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYTLRDLDDTSV VEDGRKKLNTI.AHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAEPKK.SH RQSIRRK.VLPEIYL.TRLLSTK.GTLQKFL.DDLF KAILSIREDK.PPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPL.FFWVNIILKNPQFVFDIDK TDHIDACL.SVIAQAFIDACSISDLQLGK.DSPTN KLLYAKEIPEYRKVQRYYKQQDMTPL.SEQE MNSHLAESRR.YQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETIQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	[ĺ	1			ACTMPEGALPAPVPVCVRFERRGCVHGNLTF
GAL SNASAPVDFFINGRAYADEVAVAEELLD PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLTUHVSTKGAGKEQDSLGQSHEY RVKIGQVSCDIQIVSDRIHCSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAHIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQRREBIRGGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFFKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAKQQNKGSIDATTG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KARLSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRIVQRYJKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							WYMQNPVITAISPRRSPVSGGRTITVAGERFH
PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HIPGEPLTLVIHVSTKGAGKEQDSIGJSHEY RVKIGQVSCDIQIVSDRJHCSVNCSLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTIFPRCSSL VEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKLLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPETYLTKLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC	1]			MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP
HPGEPLTLVIHVSTK GAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ LPTITQVGNFNQTLATLQLGSESTAIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQITDMTDLTKELNRSQ GIPFLEYKHFVTRIFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDATIG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLYTDELAEPFKSH RQSHRKKV.PETYLTRLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC			Ì				GALSNASAPVDFFINGRAYADEVAVAEELLD
HPGEPLTLVIHVSTK GAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ LPTITQVGNFNQTLATLQLGSESTAIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQITDMTDLTKELNRSQ GIPFLEYKHFVTRIFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDATIG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLYTDELAEPFKSH RQSHRKKV.PETYLTRLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	[:		ĺ				
LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQIDMTDLTKELNRSQ GIPFLEYKHFVTRITFIPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLD ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRF WVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRFQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY
LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVPCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQIDMTDLTKELNRSQ GIPFLEYKHFVTRITFIPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLD ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRF WVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRFQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	1			1		. 1	
LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQIDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASILITIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDATIG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KARLSIREDKPPLAVKYFFDFLEQAEKGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRIFFPRCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPMEE GISLFSSLLDNKHFLIVPVHALEQQKDFAVRD RCSLASSLITIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDATTG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKLEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEFKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHHWKTNSLPLFFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC							EMESOIREEIRKGFAELOTDMTDLTKELNRSO
LNSQGSSQAQETHPLLGEWKIPESCRPNMEE GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFFILVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC	1					,	
GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDITDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC]						· · · · · · · · · · · · · · · · · · ·
RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYLLRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDITDILTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKPEGASLAMSLIDKKD NTLGRVKDLDTEKYPHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNINVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVK YFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVK YFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTIGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHHWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHHWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	ļ						
NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	1						VEDGRKKLNTLAHYKIPEGASLAMSI IDKKD
RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC] !					•	
KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC							
KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	ļ l				ļ	•	
MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				[ĺ	
RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC						l	
730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC						l	
730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC		.]				ļ	
VANGKGNORNMGSPOPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC	730	2080		5744	-	202	
LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD	,,,,	2000	Α.	44/د	ا ا	474	
LKD					1	ŀ	VAINGRUNGKNINGSPUPSLLAFERNLELQIMG
731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC				' I	[
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	721	2001	_	5717	, 	202	
FRVDEVNWITWNINVGIINEDPGNCEGVKRT	131	2001	А	5/4/	1	382	
	لــــا					<u>l</u>	FKVDEVNWI I WNI NVGIINEDPGNCEGVKRT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]	peptide] -	/=possible nucleotide deletion, \=possible
				sequence	L	nucleotide insertion
						LSFSLRSSRVSGRHWKNFALVPLLREASARD
						RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS
						GEK
732	2082	A	5753	198	3	AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP
<u></u>		Ļ <u>. </u>				PGKLQAQLPCPSPVRFTSARIPPASRPQTKS
733	2083	Α	5754	2	2223	AAGPPGLEAEGRAPESAGPGPGGDAAETPGL
1 :						PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL
i l						SVANCLGAQTVQAPAEPAAGKAEQGETSGR EAPEAPAVGREDASAEDSCAEAGASGAADG
						ATAPKTEEEEEEEETAEVGRGAEAGASGAADG
						LNRTSTSTKSAKSGSEASASASKDALQAMILS
						LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN
1		ļ.				LFNINPDKGIOFLISRGFIPDTPIGVAHFLLORK
]		·				GLSRQMIGEFLGNSKKQFNRDVLDCVVDEM
1						DFSSMELDEALRKFQAHIRVQGEAQKVERLIE
	1					AFSQRYCMCNPEVVQQFHNPDTIFILAFAIILL
						NTDMYSPNIKPDRKMMLEDFIRNLRGVDDG
						ADIPRELVVGIYERIQQKELKSNEDHVTYVTK
1 1		-				VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV
						NKLQKQAAHQREVFLFNDLLVILKLCPKKKS
1						SSTYTFCKSVGLLGMQFQLFENEYYSHGITLV
						TPLSGSEKKQVLHFCALGSDEMQKFVEDLKE
						SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA
1						QGDPQSKQGSPTAKREAALRERPAESTVEVSI
1 1						HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR
1 1						QQTPPLPPPPPTPPGTLVQCQQIVKVIVLDKPC
1			1			LARMEPLLSQALSCYTSSSSDSCGSTPLGGPG
						SPVKVTHQPPLPPPPPPYNHPHQFCPPGSLLH GHRYSSGSRSLV
734	2084	A	5788	8	362	SSVMGDLVGQGLEEQIVARDENSWLIDGGTP
'37	2004	Α.	3766	8	302	IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR
]	j					KIPKRTDSVKFAGYKFEVVDIDNYRIDOLLVT
						RIDSKATALSPKLPDAKDKEESVA
735	2085	A	5827	1	1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITL
				-		PPPFQHPDLSPLLRYSFETMAPTGLSSLTVNST
						AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG
	ĺ		(NLVVCLMVYQKAAMRSAINILLASLAFADM
						LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF
						FWLFVIEGVAILLIISIDRFLIIVQRQDKLNPYR
						AKVLIAVSWATSFCVAFPLAVGNPDLQIPSRA
1	. 1	l	l			PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY
1		į				SFMGILNTLRHNALRIHSYPEGICLSQASKLGL
						MGLQRPFQMSIDMGFKTRAFTTILILFAVFIVC
]]						WAPFTTYSLVATPSKHFYYQHNFFEISTWLL
1 1		ł	ł	ł		WLCYLKSALNPLIYYWRIKKFHDACLDMMP
1 772	2000				260	KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	A	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKQFSRS
		1		ł	l	DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP
1777	2007		F071		501	LTSEVESSASGSGPGPAPSFTTCL
737	2087	A	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR
		- 1	ļ			RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH
		1			1	GLARRLTTARRPPASSEQAQQELFNELKPAV
]	J	İ	J		ļ	DGANFIVNHMRDQNNYNEEKDSWNRVART
	ĺ	Ţ			ļ	VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP QPFPGDPYSYNVQDKRFI
738	2088	Ā	5881	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG
'30	2000	^	2001	•	1100	LLDSSKLCDYENRFNTSKGGELPDRPAGVGV
			ļ			YSAMWQLALTLILKIVITIFTFGMKIPSGLFIPS
1	i	i				MAVGAIAGRLLGVGMEQLAYYHQEWTVFNS
<u></u>						Crantorano - Orthografi i i i i i i i i i i i i i i i i i i

No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	
	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
194	eotide	seq-		USSN	location	corresponding	
1914 ng to first amino aeid residue of peptide sequence 1914 ng to first amino aeid residue of peptide sequence 1915 ng to first amino aeid residue of peptide sequence 1915 ng to first amino aeid residue of peptide sequence 1915 ng to first amino aeid residue of peptide sequence 1915 ng to first amino aeid residue of peptide sequence 1915 ng to first amino aeid residue of peptide sequence 1915 ng to first amino aeid residue of peptide sequence 1915 ng to first amino aeid residue insertion 1915 ng to first amino aeid residue aeid aeid aeid aeid aeid aeid aeid aei	seq-		[
mino acid residue of peptide requence per per per per per per per per per pe			1	1		1	
Peptide Pept	1 20.1.00		1	/ **			
Peptide	1			1			1-Threonine, v=vanne, w=tryptopnan,
		Į.		i		sequence	Y=1 yrosine, X=Unknown, *=Stop codon,
WCSQGADCTPGLYAM/CAAACIGGYTEMT		Í	İ	}			
VSLVVIMFELTGGLEYTYPIMAAAMTSKWVA DALGEGEGYDAHRINGPYPIEAKEEPHKTL AMDVMKPRNDPILITVLTQDSMTVEDVETII SETTYSEPTVVSRESQRULGYPUREPHANDLISIE NARKKQDQVVSTSITYFTEHSPPLPPYTPPTLK LENILLDSFTYTVDLTPMELVVDIPFKLIGIRQC LVTHNGRLIGITKKDVLKHAQMANQDPDSI LEN LEN LDSFTYTVDLTPMELVVDIPFKLIGIRQC LVTHNGRLIGITKKDVLKHAQMANQDPDSI LEN LEN LDSFTYTVDLTPMELVVDIPFKLIGIRQC LVTHNGRLIGITKKDVLKHAQMANQDPDSI LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LEN LOSA LEN LEN LEN LOSA LEN LEN LEN LEN LOSA LEN					sequence	<u> </u>	
VSLVVIMFELTGGLEYTYPIMAAAMTSKWVA DALGEGEGYDAHRINGPYPIEAKEEPHKTL AMDVMKPRNDPILITVLTQDSMTVEDVETII SETTYSEPTVVSRESQRULGYPUREPHANDLISIE NARKKQDQVVSTSITYFTEHSPPLPPYTPPTLK LENILLDSFTYTVDLTPMELVVDIPFKLIGIRQC LVTHNGRLIGITKKDVLKHAQMANQDPDSI LEN LEN LDSFTYTVDLTPMELVVDIPFKLIGIRQC LVTHNGRLIGITKKDVLKHAQMANQDPDSI LEN LEN LDSFTYTVDLTPMELVVDIPFKLIGIRQC LVTHNGRLIGITKKDVLKHAQMANQDPDSI LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LOSA LEN LEN LEN LOSA LEN LEN LEN LOSA LEN LEN LEN LEN LOSA LEN	1	ł		1			WCSQGADCITPGLYAMVGAAACLGGVTRMT
DALGREGIYDAHRINGYPIEAKEEPHAHL AMDVMKPRRNDPLITVITODSNITVEDVETII SETTYSGFPVVVSRESQRLVGFVLRRDLISSE NARKKQDGVVSTSINYTEISPEPYPPTPITLK LRNII,DLSFTVTDLTPMEIVVDIFRKLGLRQC LVTHNGRILGIITKKDVLKHIAQMANOPDSI LEN LRNII,DLSFTVTDLTPMEIVVDIFRKLGLRQC LVTHNGRILGIITKKDVLKHIAQMANOPDSI LEN T.QLAASVPFBAISLISWVLPESARWLIINGKP DQALQELRKVARINGHKEAKNLTEVILMSS KEEVASAKEPRSVDLFCVPVLWRSCAMI V VNFSLLISVYGLVFDLGSIGRDIFLQALFGG VPLGRATTALLISLIGRETIQAGSQAMAGI AILANMI VPDLCTLRVVFAVLGKOGFGISL TCLTYKAELFPTFVRMT ADGIHTIVGRLGA VPLGRATTALLISLIGRETIQAGSQAMAGI AILANMI VPDLCTLRVVFAVLGKOGFGISL TCLTYKAELFPTFVRMT ADGIHTIVGRLGA VPLGRATTALLISLIGRETIQAGSQAMAGI AILANMI VPDLCTLRVFAVILGKOGFGISL TCLTYKAELFPTFVRMT ADGIHTIVGRLGA VPLGRATTALLISLIGRETIQAGSQAMAGI AILANMI VPDLCTLRVFAVILGKOGFGISL TCLTYKAELFPTFVRMT ADGIHTIVGRLGA VPLGRATTALLISLIGKRT VPLGRATTALISLIGKT VPLGRATTALISLIGKRT VPL		[1		İ	VSLVVIMFELTGGLEYTVPLMAAAMTSKWVA
AMDVMKPRNDPLLTVLTQDSMTVEDUTE SETTYSGFPVVVSRSQRLVGPVTRDLSISE NARKKODGVSTSIIYTTEHSPPLPPYTPPTILSE NARKKODGVSTSIIYTTEHSPPLPPYTPPTILSE NARKKODGVSTSIIYTTEHSPPLPPYTPPTILSE NARKKODGVSTSIIYTTEHSPPLPPYTPPTILSE NARKKODGVSTSIIYTTEHSPPLPPYTPPTILSE LYTHNGRILGIITKKDVLKHIAQMANQDPDSI	ŀ		ļ				
			i	1		-	
NARKKODGVSTSITYTEHSPFLPPYTPPTILL	1			l			
LRNILDISPITYTDLITMEIVYDIPRKIGIRQC		•					
LVTHNGRLIGITIKKDVLKHIAQMANQDPDS	1		ŀ	1 .		ļ	
1739 2089 A 5892 2 916	1			i i			
739	1					'	
DQALQELRKVARINGHKEAKNITIEVLMSSV KEEVAAKEPRSVLDIFCVPVLRWRSCAMLV VINPSLLISYYGLVPLQSLGGEGEAMLV VINPSLLISYYGLVPLQSLGGEGEAMLO ALAMMLVPQLQTLRVVALGKGCFGISL TCLTYKAELFPTPVRMTADGILHTVGRIGA MMGPILIMSRQALPLIPPLYGVISIASSLVVL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEVALHGAL 740 2090 A 5900 2 426 RPIKTLGIGFHFSVDGVVFLTQREVQNLWKE NLILIDTAKKHGYVEVUDTFITTMGRYKEELQG KCGCHHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPRQV CSEILLSRMCANRITM GRYFSNQSKLQQGTVTNFRSPYHVRGPRQV CSEILLSRMCANRITM GRYFSNQSKLQQGTVTNFRSPYHVRGPRQV CSEILLSRMCANRITM GRYFSNQSKLQQGTVTNFRSPYHVRGPRQV CSEILLSRMCANRITM GRYFSNQSKLQQGTVTNFRSPYHVRGPRQV CSEILLSRMCANRITM GRYFSNQSKLQQGTVTNFRSPYHVRGPRQV CSEILLSRMCANRITM GRYFSNGSKLQQGTVTNFRSPYHVRGPRQV CSEILLSRMCANRITM GRYFSNGSKLQGGTVERFYBYLGGTYSEPG KFWV VSYEIKDMCKGCHFISVKSKLRLEISEKRER QRELKEKIREERRINLAABMGBEDGEKEPQEE EEEKEEEEEEEEPLPEIFPSTPSPLCGFYSEPG KFWV TQMGNDKSIKCEQNLGHDTMYWYRQDSKK FLKIMFSYNNKELIINETVPNRSPSKSPDFACGFYSEPG KFWV TQMGNDKSIKCEQNLGHDTMYWYRQDSKK FLKIMFSYNNKELIINETVPNRSPSKSPDFALQSHCTVHKPPGSARKLQGSVCTTQGSSLHSILMASDO VPVC VFVC HKPPGSARKLQGSVCTTQGSSLHSILMASDO VPVC TQMGNTASSWSRSRDTALQSHCTVH TGMGNTAVSQTPKYLV TGMGNTAVSTAVELIINETVPNRSPSKSPDFALQSHCH GLGRQTAKSSWSRSRDTALQSHCTVH TGMGNTAVSTAVELIINETVNNN DTKHHLYSDILGDGGAWGGSLEWSRQSADRRAL GLSRQTAKSSWSRSRDTTCCCRRAWMLVPA ADRARERFINNEK WDTNSSENWHPIWNN DTKHHLYSDILGUFGMPTILDMINVNN THALASDLVQRGGFAWGGSLEWSRQSADRRAL GLSRQTAKSSWSRSRDTTCCCRRAWMLVPA ADRARERFINNEK WDTNSSENWHPIWNN THALASDLVQRGGFAWGGSLEWSRQSADRRAL GLSRQTAASSWSTLVAILADDRFQCVVY PFKPKLTIKTAFANINLABSLAUMPGRGGTM CKISGLVQGISVAASVFILVAILADDRFQCVVY PFKPKLTIKTAFANINLABSLAUMPGRGGTM CKISGLVQGISVAASVFILVAILADDRFQCVVY PFKPKLTIKTAFANINLABSLAUMPGRGGTM TRIQUVQSTGPROPHOTELTHYRGRTTMSPFRANCH TRIQUVQSTGPROPHOTELTHYRGRTTMSPFRANCH TRIQUVQSTGPROPHOTELTHYRGRTTMSPFRANCH TRIQUVQSTGPROPHOTELTHYRGRTTMSPFRANCH TRIQUVQSTGPROPHOTELTHYRGRTTMSPFRANCH TRIQUVQSTGPROPHOTELTHYRGRTTMSPFRANCH TRIQUVQSTGPROPHOTELT							
DQALQELRKYARINGHKEAKNILTEVIMSSY KEEVASAKEPSVIJLIFCPVIRSCAMILY VNFSLISYYGLYFDLQSLGRDIFILQALFGA VDFIGRATTALLLSFLGRRTIQAGSQAMAGL ALAMMLYPQDLQTLRVVPAVIGKGCFGISL TCLTTYKAELFPTPVRMTADGILHTYGRIGA MMGPILLMSRQALPLIPPLIYQVISIASSLVVL FTLPETGQLFLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEIPVALHGAL 740 2090 A 5900 2 426 RPIKTLGIGFHFSVDGVVFILTQREVQNLWE MILIDTAKKHGVEVVDTFTTMGRYKEFLQG KCGCHHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTMFRSPYHVRGPRQV CSEILLSRMCANKRTM GRYFSNQSKLQQGTVTMFRSPYHVRGPRQV CSEILLSRMCANKRTM GRYFSNQSKLQQGTVTMFRSPYHVRGPRQV CSEILLSRMCANKRTM GRYFSNQSKLQGGTVTMFRSPYHVRGPRQV VSYEIKDMCKGCFHSSVKSKLIRLEIERKER QRELKEKIREERINKLAABMGEDGEKEFQEE EEEKEEEEEEEPLFEIFFSTPSPLCGFYSEPG KFWV GRELKEKIREERINKLAABMGEDGEKEFQEE EEEKEEEEEEEPLFEIFFSTPSPLCGFYSEPG KFWV GRELKEKIREERINKLAABMGEDGEKEFQEE EEEKEEEEEEEPLFEIFFSTPSPLCGFYSEPG KFWV GRELKEKIREERINKLAABMGEDGEKEFQEE EEEKEEEEEEEPLFEIFFSTPSPLCGFYSEPG KFWV GRELKEKIREERINKLAABMGEDGEKEFQEE EEEKEEEEEEEPLFEIFFSTPSPLCGFYSEPG KFWV GRELKEKIREERINKLAABMGEDGEKEFQEE EEEKEEEEEEEPLFEIFFSTPSPLCGFYSEPG KFWV GREKKIREELINETVPNFSFKSPDKAHL HINHISLEIGBAVYFCASSDTALQSBCTPKLA HINHISLEIGBAVYFCASSDTALQSBCTPKLA GREKKEEPT GREKK	739	2089	Α	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP
KEEVASAKEPRSVLDLGVCPVLRVRSCAMLV VNFSLLISYVGLVFDLGSLGDFILQAJEFOA VDFLGRATTALLLSFLGRRTIQAGSQAMAGL AILANMLVPQDLQTLRVVFAVLGKGCFGISL TCLTIYKAELFPTPVRMTADGILHTVGRLGA MMGPLILMSRQALPLPPLLYGVISASSLVVL FFLPETGGLPPDTIQDLSSQKSTAAQGTRQE AFTVESTSLLEVALHGAL 740 2090 A 5900 2 426 RPIKTLGGFISVDGVPHLTQREVQNLWKE NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG KCGCHFIEVVXSKLSKEYNFIKMFSRNHIM GRYFSNSKLQQGTVTNFRSYHVRGPRIQV CSEILLSRMCANKRTM RPMFSTLLICHNGYILEAPLPTIKQEEDDHDV VSYSIKDMCIKCFHFSSVSKSLLRIEBKAER QSEKEEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV KRIMPSTRILLICHNGYILEAPLPTIKQEEDDHDV VSYSIKDMCIKCFHFSSVSKSLLRIEBKAER QSEKEEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV KRIMPSTRILLICHNGYILEAPLPTIKQEEDDHDV VSYSIKDMCIKCFHFSSVSKSLLRIEBKAER QSEKEEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV KRIMPSTRINGERINGTPVFNRFSPKSPDKAHL NLHINSLEIGDSAVYFCASSQDTALQSHCTPV KRYPGSARKLQGSVCTCTQGSSLHSLMASDO VPVC KRYPGSARKLQGSVCTCTQGSSLHSLMASDO VPVC KRYPGSARKLQGSVCTCTQGSSLHSLMASDO VPVC KRYPGSARKLQGSVCTCTQGSSLHSLMASDO VPVC A S938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER ALSVQQRGGPAWSGSLEWSRQSAGDRRL GJSRQTAKSSWSRSRDMFTCCCRRAWMILVPA ADRARRERFINMEKWDTYNTVNYLQFQCVVY PFKPRLTIKTAFVIMIMIVATHISFBAVMLH VQEKYYRVRLNSQNKTSPYVVCRSWPNQ EMKRYTTIVLANIYLAPILTJSLVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIKMLLU VALLFILSWLPWTLMMISTANISPSAVMLH VQEKYYRVRLNSQNKTSPYVVCRSWPNQ EMKRYTTIVLANIYLAPILYJSLVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIKMLLU VALLFILSWLPWTLMMISTANISPSAVMLH VQEKYYRVRLNSQNKTSPYVVCRSWPNQ EMKRYTTIVLANIYLAPILYJSLVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIKMLLU VALLFILSWLPWTLMMISTANISPSAVMLH VQEKYYRVRLNSQNKTSPYVVCRSWPNQ EMKRYTTIVLANIYLAPILTJKTNSPSAVMLH VQEKYYRVRLNSGNCTSPYVVCRSWPNQ EMKRYTTIVLANIYLAPILTJKTNSPSAVMLH VQEKYYRVRLNSGNCTSPYVVCRSWPNQ EMKRYTTIVLANIYLAPILTJKTNSPSAVMLH VQEKYYRVRLNSGNCTSPYVVCRSWPNQ LVMEELKETTNSSEI SHVCVSHTAGSSGCFAGGGTAGAAALGISAVA LVDYQGGRLGVARGAWYMEAPDRQGDM LVMEELKETTNSSEI SHVCVSHTAGSSGCFAGGGTAGS LASPQCVCCGALWWMLLVGLGFFFIMEALSIFLYVPTL GOVCCGALWWMLLVGLGFFFIME	J		J			1	
VNFSLLISYYGL.VFDLQSLGRDIFLQALFGA VDFLGRATTALLIS-ILGRATIQAGSQAMAGI. AILANMLVPQDLQTLRVVFAVLGKGCFGISI. TCLTIYKAELFPTPVRNTADGILHTVGRIGA MIMGPLILMSRQA.PLLPFLYGVISLASSL.VVL FFLPETQGI.PLPDTIQDLESQKSTAAQGTNRQE AFTVETSILEPVALHGAI.			l				
VDFLGRATTALLLIST-GRRTIOAGSQAMAGL ALLANML-VPQLQTLR-VVPAVLGRGCFGISL TCLTTYK-ABLPPTPVRNTADGILHTVGRLGA MMGPLLMSRQALPLPPLLYGVISGASSL-VVL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEIVALHGAL 740 2090 A 5900 2 426 RPIKTLGGFHSVDOVHFLTGREVQNLWKE NLIILDTAKKHGYEV-VDTFTITMGRYKEFLQG RCGCHFHEVVSKSLSKYSKYFKKMRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM CSEILSRMCANKRTM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM CSEILSRMCANKRTM CS			l				VNEST LISVYGI VEDI OSI GRDIELI OALEGA
AILANML.VPQDLQTLR.VVFA.VLQRGCGFGISL TCLTIYK.AEL.PFPVMTADGI.HTVGRLQA MMGPLILMSRQALPLLPPLLYGVISLASSLVVL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEVALIGAL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEVALIGAL RPIKIT.GIGPHPSVDGVHPLTQREVQNLWKE NLIILDTAKKHGYEVDTITTMGRYKEPLQQ KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSKMCANKRTM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSKMCANKRTM CSE			1				
TCLTTYKAELFPTPVRMTADGILHTVGRLGA	1		1				
MMOPLILMSROALPILIPPILIYGNISIASSLVYIL FFIPETOGIPLPPTIDIGENGKSTAAQGINRQE	1		1				
FFLPETQGLPLPDTIQDLESQUSTAAQGNRQE			1				ICLITYKAELFPIPVKMTADGILHTVGRLGA
AATVESTSLEIPALHGAL	1		1				MMGPLILMSRQALPLLPPLLYGVISIASSLVVL
740							
NLIILDTAKKHGYEVVDIFTITIMGRYKEFLQG KCGCHHEVVKSKLSKEYNFIKMKRSRNHM GRYFSNQSKLQQGTVINFRSPYHVRGPINQV CSEILLSRMCANKRIM 741 2091 A 5910 3 412 RMPESTILLICENGYILEAPLPTIKQEEDDHDV VSYEIKDMCIKCFHFSSVKSKLIRLIEEKRER QRELKEKRERRINKLAAEMGEDGEKEFQEE EEKEEEEEEEPLPEIFIPSTPSPLCGFYSEPG KFWV 742 2092 A 5936 1 482 MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEGNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETYPNRSPKSSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRL GLSRQTAKSSWSRSDRTCCCRAWWILVPA ADRARRERTIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHOPQVAAIFIISYP LIFFLCMMGNTVVCFIVMRNKHMITVTNLFI LNLAISDLLVGIFCMPITLLDNILAGWFFGNTM CKISGLVQGISVAASVFILVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIWVALATIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANTYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVPIPIYGFFNENFRRG FQEAFQLQLQCKRAKPMEAYALKAKSHIVLIN TSNQLVQESTFQNPHGETILLYRKSAEKPQQE LVMEELKETTNSSI TSNQLVQESTFQNPHGETILLYRKSAEKPQQE LVMEELKETTNSSI TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSI TSNQLVQGSGCPAGAGAGAVALGISAVA LYDVQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAFTPMSGLGSGRGROGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV CGALMWIMLILVGLGFFPDMARPAACMV			<u> </u>				
NI.III.DTAKKHGYEVVDTTTITMGRYKEFLQG KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTIM RMFPSTILLICENGYIL.EAPLPTIKQEEDDHDV VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER QRELKEKIREERRINKLAAEMGEDGEKEFQEE EEEKEEEEEEEPLPEIFPSTFSPILCGFYSEPG KFWV 742 2092 A 5936 1 482 MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLEI.GDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRIL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAHISYF LIFFLCMMGRIVCFIVMRNKHMITVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFILVAIAVDRFQCVVY PFKFKLTIKTAFVIIMITWVLAITIMSPSAVMLH VQEEXYYRVRLINSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQCWHVVSKKQKIKMLLI VALLFILSWLPLWTLMMLSDVADLSPNELQII NIYYTPAHWLARGQCWHVVSKKQKIKMLLI VALLFILSWLPLWTLMMLSDVADLSPNELQII NIYYTPAHWLARGQCWHTVSSKKQKIKMLLI VALLFILSWLPLWTLMMLSDVADLSPNELQII NIYYTPAHWLARGGGMGGTAKAKSHVLIN TSNQLVQESTFQNPHGETILYRKSAEKPQQE LVMEELKETTINSPSAVMLI TSNQLVQESTFQNPHGETILYRKSAEKPQQE LVMEELKETTINSGNSVPPIIVGFPNENFRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETILYRKSAEKPQQE LVMEELKETTINSGRGGGTAGGAGAVALGISAVA LYDYQGGRIGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTOWAWAPTPMSGLGGGRGGGTLASS PLSLPLLLAGVTGLATELFOQMARPAACMV CGALMWIMLILVGLGFFPIMEALSHFLYVFFL GVCVCGALYTGLILTELFOQMARPAACMV	740	2090	Α	5900	2	426	RPIKTLGIGFHFSVDGVHFLTOREVONLWKE
KCGCHFIEVYSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM RMPESTILLICENGYILEAPLPTIKQEEDDHDV							NLIILDTAKKHGYEVVDTETTTMGRYKEFLOG
GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKTRY CSEILLSRMCANKTR	1						
CSEILLSRMCANKRTM	1						
741							
VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER QRELKEKIREERRINKLAAEMGEDGEKEFQEE EEEKEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV	741	2001		5010	-		
QRELKEKIREERRNKLAAEMGEDGEKEFQEE EEEKEEEEEEEPPLPEIFIPSTPSPIL.GFYSEPG KFWV 742 2092 A 5936 1 482 MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRRL GLSRQTAKSSWSRSRDRTCCCRRAWHLVPA ADRARRERIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHOPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVINLFI LNALSIDLUVGIFCMPITLLDNILAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPILSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VAALFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFONPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDPYQGGRIGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLFLLLAGVTGILATELFDQMARPAACMV CGALMWIMLI VGLGFPFIMEALSHFLYVFFL GVCVCGALMYMGLI VGLGFPFIMEALSHFLYVFFL	/41	2091	A	2910	3	412	
742 2092 A 5936 1 482 MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLEIGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYDINITYVYLHQPQVAAIPIISYP LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNILAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRPQCVVY PFKPKLTIKTAFVIIMIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVIPIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWILLIVGLGFPFIMEALSHFLYVPFL GVCVCGALYTGLLFDETKGKTFQEISKELHRL							
742 2092 A 5936 1 482 MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIPIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNILAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRPQCVVY PFKPKLTIKTAFVIIMIIWVLATIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIKMLLI VALLFILSWLPLWTLMMLSDVADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDVQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLII.VGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	1						QRELKEKIREERRNKLAAEMGEDGEKEFQEE
742 2092 A 5936 1 482 MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRL GLSRQTAKSSWSRSRDRTCCCRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNITVVCFIVMRNKHMHTVTNLFI LNIAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFILVALAVDRPQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVILNSQNKTSPYYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTPFAHVALAFGNSSVNPIYGFFPENFRGG FQEAFQLQLCCKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM ALYDYQGGRLGVARGAWYMEAPDIRQGDM GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWILLI CLGFPFIMEALSHFL,YVPFL GVCVCGALYTGLFPETKGKTFQEISKELHRL	1 1						EEEKEEEEEEEPLPEIFIPSTPSPILCGFYSEPG
TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIPIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNILAGWPFGNTM CKISGLVQGISVAASVFILVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEKYYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NITYPPFAHWLAFGNSVNPITYGFFNEPRRG FQEAFQLQLCQKRAKPMEAYALKAKSHYLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTINSEI 744 2094 A 5966 149 327 SHCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGLAFFIMEALSHFLLYVPFL GVCVCGALYTGLFPFIMGALSHFLLYVPFL GVCVCGAIYTGLFPEKGKTFQEISKELHRL	i						KFWV
TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHNSLEIGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQV AAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIAGWPFGNTM CKISGLVQGISVAASVFILVALAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVNPIIVGFFNENFRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTINSEI 744 2094 A 5966 149 327 SHCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGLAFFIMEALSHFLYVPFL GVCVCGALYTGLFPFIMEALSHFLYVPFL	742	2092	Α	5936	1	482	MGCRLLCCVVFCLLOAGPLDTAVSOTPKYLV
FLKIMFSYNNKELINETVPNRFSPKSPDKAHL NLHINSLEIGBSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 I 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRIL GLSRQTAKSSWSRSRDRICCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYP LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITILDNIIAGWPFGNTM CKISGLVQGISVAASVFILVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFFFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL))	ı j		}			
NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRPQCVVY PFKKPLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVIPITYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL							
HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC 743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNILAISDLLVGIFCMPITLLDNILAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	1 1				1		
743 2093 A 5938 1 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRIL GLSRQTAKSSWSRSRDRTCCCRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNILAGWPFGNTM CKISGLVQGISVAASVFILVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPYYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGA WYMEAPDIRQGDM THE COMMENT OF THE				•	l		HUDDOG A DVI COMICTOTO CONTINUE CONTINU
743 2093 A 5938 I 1566 MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIPIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTI.VAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVNPITYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM LYDYQGGRLGVARGAWYMEAPDIRQGDM GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPILLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFFFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	, 1)	J		
RALSVQQRGGPAWSGSLEWSRQSAGDRRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHILYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTINLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRPQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIVTYPFAHWLAFGNSSVNPITYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFPETKGKTFQEISKELHRL	1	2005		7005			
GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIIFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFILVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFFFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	743	2093	A	5938	I	1566	
GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIIFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFILVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFFFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL]			ľ			RALSVQQRGGPAWSGSLEWSRQSAGDRRRL
ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANITYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	[GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA
DTKHIHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLFLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL				1		l	ADRARRERFIMNEK WDTNSSENWHPIWNVN
LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL] [1	ſ	į		ĺ	TOTAL TION TO THE REAL PROPERTY OF THE PARTY
LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYTYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	1		. [ļ		ļ	
CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFFFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL		ļ	I		į	ļ	
PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFFFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	1 1	1	,	ļ	J		
VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	ļ	1			ļ		
EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	1 1	İ			ı		
RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	į l		i				VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ
RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL		ļ			1		EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF
VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLFLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL		ł]		ĺ		
NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	j i	ľ	1	i	- 1	l	
FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI 744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	[[.	j		1	Ī	NIVIVPEAHWI AEGNSSIADIIVGEENENEDDG
744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	j		i				
744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	}	Ì		1	l		
744 2094 A 5966 149 327 SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM 745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	; 1	İ	ſ	ĺ	- 1	ĺ	• • • • • • • • • • • • • • • • • • • •
745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	 	2001					
745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	744	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
745 2095 A 5970 413 856 GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	L					1	LYDYQGGRLGVARGAWYMEAPDIRQGDM
PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	745	2095	A	5970	413	856	
CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	1	- {	. 1		·	1	
GVCVCGAIYTGLFLPETKGKTFQEISKELHRL	ļ .i				1	1	
			- 1	Ī	l		
NFPRRAQGPTWRSLEVIQSTEL							•
	<u> </u>						NFFKKAQGF1 WKSLEVIQSTEL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	neuce	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ			peptide	i -	/=possible nucleotide deletion, \=possible
1		1		sequence	}	nucleotide insertion
746	2096	Α	5971	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
1				_	*	RCARHGACORSCLASODPYCGWHSSRGCVDI
						RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
			i			SGPGDSAYGVRRDLPPASASRSVPIPLLLASV
						AAAFALGASVSGLLVSCACRRAHRRRGKDIE
						TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
1			}			VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE
						LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR
						GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL
						EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP
1						APALLGGPSPRPHECASPLRLDVPPEGRCASA
1			1			PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL
				·		
						LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG
,						RALKRVDVEKPQLSLKPPLVGPSSRQAVPNG
L	0000		5000			GRFNF
747	2097	A	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI
						LKLEQENCTLVTTFRGHTGGVTALCWDPVQ
				,		RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN
1						DRVQALSYAQHTRQLISCGGDGGIVVWNMD
						VERQETPEWLDSDSCQKCDQPFFWNFKQMW
						DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL
						MGFEFEVRVCDSCHEAITDEERAPTATFHDSK
4 1						HNIVHVHFDATRGWLLTSGTDKVIKLWDMT
						PVVS
748	2098	Α	6001	2 .	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV
						CLVLLVANILRILFWFGRRFESPLLWQSAIMIL
						TMLLMLKLCTEVRVANELNARRRSFTAADS
1						KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
1						QCVLAFTGVAGYITYLSIDSALFVETLGFLAV
						LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM
1 .						WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
						DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
749	2099	Α	6002	2	447	GRPDRSELVRMHILEETFAEPSLQATQMKLK
						RARLADDLNEKIAORPGPMELVEKNILPVDSS
						VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP
, ,						DQPASQESQGSAASPSEPKVSESPSPVTTNTP
						AOFASVSPTVPEFLKTPPTAD
750	2100	A	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG
'30	2100	А	0007	-	₹41	WRWELRLRNYVPEDEDLNKRRVPQAKPDAV
						OPEN WEGGE EAAR CONDON WEEDING AREA ON CONDO
'						WDLKRDVAKKLEKLLKRTQRAIAELIRERLK
				ļ		
751	2101		6007	22	1200	GQEDSLDSAVDAATEHKTC
/31	2101	A	6007	33	1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF
[[İ		i		SHPDKLKRMSKSVPAFLQDESDDRETDTASE
						SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS
						VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL
, ,	J			l		HVFVAQCKDLAAADVKKQRSDPYVKAYLLP
			ļ			DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK
				ŀ		QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE
	•			ļ		TWDWDNKQNKQLRWYPLKRKTAPVALEAE
1 1		i		•		NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
		ļ			İ	KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ
		1				KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC
		1				VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT
						EVDWMDSTSEEVALWEKMVNSPNTWIEATL
1	- 1	ł		1	i.	PLRMLLIAKISK
752	2102	Α	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL
ļ ·						SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF
			İ			AAAIPGHRCWVHMLDNNTGSGNETGILSEDA
						

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG TIHSTSEADTEPCVDGWVYDQSYFPSTIVTKW DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV
						SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL VVILSSGALNIGQIILGGLAYVFRDWQTLHVV ASVPFFVFFLLSRWLVESARWLIITNKLDEGL KALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICILVFLRK KNLKEKA
753	2103	A	6043		1470	DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ SQAEIRKQILGSSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD YFLNCSALRMADVIRAEILEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNL PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ
754	2104	A	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF LMLINVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVR LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS
755	2105	A	6059	3	1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE RNARSVKITKRFTKLLIAPESAAPEEALGPAEE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL QGPAGIGKTMAAKKLYDWAAGKLYQGQVD FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAAP CTDPFEAASGARVLGGLLSKALLPTALLLVTT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLG GG
756	2106		6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS ARDLIMNNLTELQPGLFHHLRFLEELRLSGNH LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRH LWLDDNALTEIPS
757	2107	Α	6063	54	419 .	ITPLGLGAADMCAFPWLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENII WSSHKSLATVVPGKEGHPATIMVTNPHYQG

SEQ ID NO: of nucl-cotide sequence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QILTMLLRSLQQPSASWPRDCSSSCSW IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC
759	2109	A	6072	3	650	LRAVLKLMSECWAHNPASRLTALRIKKTLAK MVESQDVKI PGRRFRPAALEERAMEKLREKVPFQNRGKGT LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW
						SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV HSSVADMQNMPAAVHALLTQPSLSAAPFAQ RYLGTLPSTGSTTLPQCHAGNATVW
760	2110	A	6077	3	730	PLRLTLMEEVLLLGLKDREGYTSFWNDCISSG LRGCMLIELPLRGRLQLEACGMRRKSLLTRK VICKSDAPTGDVLLDEALKHVKETQPPETVQ NWIELLSGETWNPLKLHYQLRNVRERLAKNL VEKGVLTTEKQNFLLFDMTTHPLTNNIKQR LIKKVQEAVLDKWVNDPHRMDRRLLALIYL AHASDVLENAFAPLLDEQYDLATKRVRQLLD LDPEVECLKANTNEVLWAVVAAFTK
761	2111	A	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRDAIRSNNINTLFKIMRVGTASS
762	2112	A	6079	2	1558	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI QLAKLDSEAWPGVLDSERDRLILINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LKGSTIIRSKTTSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRPKLPDNEVHWGSII IQASTMIISRVPNISVHLLHEPPALTNEMYCLV

	Fono w	T	1.000	r Baller Trans	F-6-10	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
cotide seq-	seq- uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	June		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide		/=possible nucleotide deletion, \=possible
1	1			sequence		nucleotide insertion
	<u> </u>					VTVQSHEKTQIRDVKLTAGLKPGQDANLTQK
						THYTLHGTELCDESYPALLTDIPVGDLHPGEQ
	!			l .		LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE
	i					KEIVCKCHKDETVTIETVFPFDVAVKFVSTKF
						EHLERVYADIPFLLMTDLLSASPWALTIVSSE
		ł		 		LHLAPSMTTVDQLESQVDNVILQTGESASECF
1				1		CLQCPSLGNIEGGVATGHYIISWKRTSAMENI
		l	1	ļ]	PITTVITLPHVIVENIPLHVNADLPSFGRVRES
		l				LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG
	1	1				LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS
						LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM
764	2114	A	6093	1	1422	DDTSIAAA
/04	2114	[^	6693	l '	1444	AAADLANSNAGAAVGRKAGPRSPPSAPAPAP PPPAPAPPTLGNNHQESPGWRCCRPTLRERN
	1					ALMFNNELMADVHFVVGPPGATRTVPAHKY
1		1				VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA
		1	[AFLILLKYMYSDEIDLEADTVLATLYAAKKYI
						VPALAKACVNFLETSLEAKNACVLLSOSRLF
						EEPELTQRCWEVIDAQAEMALRSEGFCEIDR
		1				QTLEIIVTREALNTKEAVVFEAVLNWAEAEC
		i	i i			KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE
						FANGAAQSDILTLEETHSIFLWYTATNKPRLD
				Ì		FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG
						RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV
]	j	ļ			KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF
		1			·	EHPVQVEQDTFYTASAVLDGSELSYFGQEGM
]	ł				TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE
765	2115	 	6000	1	1150	LIFYA
765	2115	A	6099	1	1150	SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF
		ſ	[RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG
			'		•	HADTCHFDVNVWEASGNRSGGVCDDCQHN
1	1	l	1			TEGOYCORCKPGFYRDLRRPFSAPDACKPCS
1		1	[CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA
		l				GRRCDRCMVGYWGFGDYGCRPCDCAGSCD
		1	1, 1			PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP
i		1]			WEWEDAQGFSALLHSGKCECKEQTLGNAKA
	Ī	}				FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK
						KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL
	•					NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK
		L				PSLGRKVMDILKRECK
766	2116	Α	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR
			1			CVLTERGLQLFEAKGTGGRPKELSFARIKAVE
					•	CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW
						NAQITLGLVKFKNQQAIQTVRARQSLGTGTL
767	0117	_	(10)			V\$
767	2117	A	6106	1	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP
						MARYYIIKYADQKALYTRDGQLLVGDPVAD
						NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC
} · ·						LACVETEEGPSLQLEDVNIEELYKGGEEATRF
1		1				TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW
768	2118	A	6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF
/00	2115	Α.	לטוס	3	292	YTPSVISSVMHRVARCAAPHVHILLANFYLLF
			l J			PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
769	2119	A·	6110	1	711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSS
'0'	2117	Δ.	0110	•	,11	SSSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ
						ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY
						KHEDLQTDESSMDDRHPRRQLCGGNQAATE
		Ц				בוצאיטיטיטאיט וואונעוייייייי אינייייייייי

L'ano III	CEO 10	F-3-7-A	Leco	I D 12 -4 - 12	1 8 30 30	
SEQ ID	SEQ ID	Met	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	in NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl- cotide	peptide		USSN	location	location corresponding	F=Phenylatanine, G=Grycine, H=Phistidine, I=Isoleucine, K=Lysine, L=Leucine,
,	seq- uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uance		914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
delice	İ		314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Í	[[İ	peptide	Sequence	/=possible nucleotide deletion, \=possible
ļ		1		sequence		nucleotide insertion
			├──	Sequence		RIILFGRELQALSEQLGREYGKNLAHTEMLQD
			1			AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
						NSAILESONLPKOPPLMLALGOASECLRLMA
		l				RAGLGSCSFARVDDYLH
770	2120	A	6125	2	570	YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
1 ''0	2120	^	0125	4	1 370	VAPWALKYMNRRASQMLLMFLLAICLLAIIF
		ļ				VPQEMQMLREVLATLGLGASALANTLAFAH
		Ì	1			GNEVIPTURARAMGINATFANIAGALAPLMM
					ŀ	ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
1		ľ	ľ		Í	PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	A	6126	909	353	RSFVLDTASAICNYNAHYKNHPKYWCRGYF
1 ′′′	2121	^	0120	303	333	
J	}	J				RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC LTKEDTGWYWCGIQRDFARDDMDFTELIVT
]					DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
	İ	[RKADRSRTSILIICILITGLGIISVISHLTKRRRS
1		ł			}	
772	2122	A	6148	7	810	QRNRRVGNTLKPFSRVLTPKEMAPTEQM
1 ′′2	2122	A	0148	′	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF
						TOGSGENKEEINYEFDTKDLVCLGLSSIVGV
	}	1				WYLLRKHWIANNLFGLAFSLNGVELLHLNN
	1					VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
		1	Ì			FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
1]					VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF
1						GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV
						ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
773	2123	A	6161	3	1088	EASASKGLEKKEK COPMLVTRKNHPKLLLRRTESVAEKMLTNW
1/3	2123	Ι Α	0101	3	1088	FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG
1						
ł	-	1				PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV NPENENAPEVPVKGLDCDTGTQAKEKLLDA
						AYKGVPYSQRPKAADMDLEWRQGRMARIIL
	! 	ł	1	•		QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV
1						ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
						SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
	•	ļ				LDQREGDRGSKMVSEIYLTRLLATKGTLQKF
1		(VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ
		İ				ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
774	2124	A	6163	860	125	NPQFVFDIHKNSITDACLSVV KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
117	2124	^	0103	300	143	SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE
		ļ				PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA
		l				GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
1.		l				DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
						NHTMVYDGFGPADLRQACAELSLWDHGALA
						NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK
						QLWQALLEQPCEWVDGLLPLRTNLAPRT
775	2125	A	6191	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
''3	لغنغ	^	0171	-	374	YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
			1	•		DMK\KSPEIISRRMTFAL*CYSLTFVRFAHYVO
			j l			\PWNWLMLGCHTAVDFDQLISSMPCISHGMT
-						ASASAL
776	2126	A	6217	1	827	FRGYWGVREAFTDASWSGGLGPGKPGMKIT
1 / 10	2120	^	021/	1	04.7	
' '			, ,			RQKHAKKHLGFFRNNFGVREPYQILLDGTFC
						QAALRGRIQLREQLPRYLMGETQLCTTRCVL
			[•	KELETLGKDLYGAKLIAQKCQVRNCPHFKNA VSGSECLLSMVEEGNPHHYFVATQDQNLSVK
			1 1		l	VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA VESG\RLSQCMRKKVSNISKRNRV**KTLNRG
						RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
] .			!			KKRKRKRIRNRSNPKVLSEKONAEGE
L			<u>. </u>		.,,	MALANAMIN NEW YEAR WARDE

[000.00	I CEC TO	1 1/	Loro	D., 3: 4-3	[N 2! - 2 - 2 1	[A
SEQ ID	SEQ ID	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	100	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
	,	i		location		
cotide seq-	seq-		USSN 09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
tience	water	Į.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ucuce		i	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide		/=possible nucleotide deletion, \=possible
1		1		sequence	ļ	nucleotide insertion
777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL*
1		1				FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF
		1				YNSKISPALWGPPVIPSALGGEAGKSL*PRRO
						RFQRGGIAPLPSRVRGRAKLFLKKK
778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN
1	"	1				AVSNADILLDLKIN*LDHILEEKMPLIEVKVVP
		1			l	PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/
					l	SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA
						LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC
1		1				NSFRYRR
779	2129	A	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP
		{				YGQSQPSCFDRVKMGFVMGCAVGMAAGAL
		1				FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM
		.				MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH
	1		L			QSQPMY
780	2130	Α	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG
1						TDYWLYSRGVCRTKSTSDNETSRKNEEVMT
1						HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE
I		l				QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV
			i 1			AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI
1		ļ				S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS
	1	1]			FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF
						LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS
						PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS
1	1	l			1	DRDHAFLQFHNSTPKEFKESLHNNPANRRTT PV
781	2131	A	6274	832	318	
/61	2131	^	02/4	032	910	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L
			}			VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS
1	1	ł				OSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD
]			LEFLGDLKGCSELKNFQELITQSALVHPKADV
	1			1		WWYCGRPLLGTLPSN
782	2132	A	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG
1 -	1	l · -				EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E
1	1] ;				DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK
						KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD
						DDKKRVKAKKKKKKKKKHKTKKKKNKKTKK
						ESSDSSCKDSEEDLSEATWMEQPNVADTMDL
1	ļ] .				IGPEAPUHTSQDEKPLKYGHALLPGEGAAMA
						EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM
	1					SGSRHRRMEAVRLRKENQIYSADEKRALASF
						NQEERRKRESKILASFREMVHKKTKGKDDK
783	2133	A	6305	201	1032	WDDYPQGALRRREAAEGLHFLGPPGRVRGQ
		,				LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP
1				l		AMARPPVPGSVVVPNWHES/RRGQGVPGLHS
						AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV
		,				SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ
		l				TTVVFWPAKLQASSRVVMFRFEFWDCGESA
1			'	ľ	·	LKKFDHMLLACMENTDAFLFLPSFTDRASFE
				l		DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT
-	L	لـــــا				DVPERDLTAFRQAWELPLLRVKSVPGRRLG
784	2134	A	6308	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP
1					i	GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA
]				ļ	PRKPEGAQARTAQSGALRDVSEELSRQLEDIL
						STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR
					ļ	TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR
			1			QSDEVGDRDHRRPQEKKKAKGLGKEITLLM
1		l J]		ļ	QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ
L	L					KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA

COEO TO	1 650 10	1 1/-4	Toro	T 8 1 4 1	18 22 2	
SEQ ID	SEQ ID	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	nou	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1 .	1	ľ	1	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}]	J	}	peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
	 	 	├	Sediferice	<u> </u>	
1				ļ		RSKLESLCRELQRHNRSLKEEGVQRAREEEE
1	1	l				KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR QENMELAERLKKLIEQYELREEHIDKVFKHK
1	ł	l	ł	1		DI COCI VIDARI COA CEL E VEA FERMIORITA
		1	ļ			DLQQQLVDAKLQQAQEMLKEAEERHQREKD
1		1]		Į	FLLKEAVESQRMCELMKQQETHLKQQLALY
1		j	ļ .			TEKFEEFQNTLSKSSEVFTTFKQEMEKMTKKI
	j .	l	ł			KKLEKETTMYRSRWESSNKALLEMAEEKTV
	1	1	1			RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR
785	2135	 	6319	1.402	000	GQRWGSHRTSAVRIFS
/83	2133	A	0319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT
1	1	ł				PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT
1				•		KLPWSWGMRPMKIFFSEEYRSISTRISHDAL*
		1				EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT
1		1				RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S
1	İ	1				QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD
	-					TTPCQKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA
		ļ				REHGQCADVDECSLAEKTCVRKNENCYNTP
		J				GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK
ì						PDTAALPRRPVMCRTYPLNYSEGCPVENVAL
		<u> </u>				RMPSPAVDSGGERLPAL
787	2137	Α	6330	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG
1	١.	i	1			SGILGLAYVMANTGVFGFSFLLLTVALLASYS
1						VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL
1						PLIEFLQSL*NSL*AVTSYEDLGLFAFGLPGKL
						VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT
		1				GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG
[ĺ					FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL
		į				NYVEKGFQISNVTDDCKPKLFHFSKESAYALP
						TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN
						TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE
						VTCHRIKDKVESELLKG***IP*SHDVVVMT\V
						KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP
		[FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV
'						GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV
H00	2.02	<u> </u>				GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	A	6351	1	6622	PRSLCFSLWAEAAVLADGGLRRRRRLLRGTM
			[SASFVPNGASLEDCHCNLFCLADLTGIKWKK
]			ļ ļ			YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV
	[LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF
						TMTYQKKKMECGRMDFPMNAVLCFSKAVH
					İ	NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN
[[ľ	KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY
					ļ	LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ
					1	AFKMSDSATKKLIGEWKQFYPISCCLKEMSE
						EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC
					l	FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS
[[İ	TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW
						VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV
		ļ '		ŀ		DRVWQECNMNRAQNKRKYSASSGGLCEEAT
		i		l		AAKVASWDFVEATQRTNCSCLRHKNLKSRN
1						AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK
				l		PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL
				l	. 1	V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE
				l		MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS
			! !	ŀ		QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ
						YQEAVEPTVYVGTAVNLEEDEANIAWKYYK
				i		FPKKKDVEFLPPQLPSDKFKDDPVGPFGOESV
1				-	ļ	TSVTELMVQCKKPLKVSDELVQQYQIKNQCL
L						SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
eotide seq-	seq- uence		09/496	correspondi	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Baice		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/-possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
						RONSEREAGKKHKVEDGTSSVTVLSHEEDA
						MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA VSYTDLDNLFNSDEDELTPGSKRSANGSDDK
						ASCKESKTGNLDPLSCISTADLHKMYPTPPSL
						EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE
						GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY
						KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE
						CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM
	·					DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI
		ļ ,				LPSPSTPRFPTPRTPRTPRTPRGAGGPASAQGS
						VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP
'						EAHSLYVNLILSESVMNLFKDCNSDSCCICVC NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV
						MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE
						KRFEALRATSAEHVNGGLKESEKLSDDLILLL
						QDQCTNLFSPFGAADQDPFPKSGVISNWVRV
						EERDCCNDCYLALEHGRQFMDNMSGGKVDE
		l.	}			ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS
						LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH
						KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV LSPFALPYWERLMLEPYGSQRDIAYVVLCPE
				_		NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR
			Ĭ			LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD
1						GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS
	:					SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS
						ATLASAASSTMTVTSGVAISTSVATANSTLTT
}						ASTSSSSSSNLNSGVSSNKLPSFPPFGSMNSNA
						AGSMSTQANTVQSGQLGGQQTSALQTAGISG
						ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL
						GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ
				:		PVKHEDREIYPOHLKSLAFSAFTOCRRPLPTS
						TNVKTLTGFGPGLAMETALRSPDRPECIRLYA
						PPFILAPVKDKQTELGETFGEAGQKYNVLFV
1						GYCLSHDQRWILASCTDLYGELLETCIINIDVP
						NRARRKKSSARKFGLQKLWEWCLGLVQMSS
						LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ
1						SLSKRLKDMCRMCGISAADSPSILSACLVAM EPOGSFVIMPDSVSTGSVFGRSTTLNMOTSOL
						NTPQDTSCTHILVFPTSASVQVASATYTTENL
					}	DLAFNPNNDGADGMGIFDLLDTGDDLDPDII
						NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD
						RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP
						LPDWFWSACPQAQYQCPLFLKASLHLHVPSV
	. •					QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN
						ALSWLTCDPATQDRRSCLPIHFVVLNQLYNFI
789	2139	A	6359	1	2002	MNML TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRR
'07	Z139 .	A	لارده	1	2002	LPVGPLLRALATCHALSRLQDTPVGDPMDLK
						MVESTGWVLEEEPAADSAFGTQVLAVMRPP
		.				LWEPQLQAMEEPPVPVSVLHRFPFSSALQRM
						SVVVAWPGATQPEAYVKGSPELVAGLCNPET
						VPTDFAQMLQSYTAAGYRVVALASKPLPSVP
						SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP
						QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA
						RGCGMVAPQEHLIIVHATHPERGQPASLEFLP
						MESPTAVNGVKDPDQAASYTVEPDPRSRHLA LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP
				_		EQKTELVCELQKLQYCVGMCGDGANDCGAL
						KAADVGISLSQAEASVVSPFTSSMASIECVPM
			<u> </u>			

C020 TO	Toro To	1 1/-4	LOFO	I m 12.4. J		TA - 1
SEQ ID NO: of	SEQ ID	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	NO: of	поа	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	peptide		USSN	location		
	seq- uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq- uence	dence	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
пенсе			314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i			peptide	sequence	
			1	sequence		nucleotide insertion
	 	├	 	Sequence		VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL
1	ł	l	1	1		YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP
ļ		1				ALVLGRVRPPGALLSVPVLSSLLLOMVLVTG
					i	VQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY
İ .						ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN
	I					NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR
		l			Í	NITDTGFKLLLVGLVTLNFVGGLHAGERARP
	1	1 .	1			VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW
			1		[PPLPAGPLR
790	2140	A	6380	76	1059	SSAGSARKLQVMALAARLWRLLPFRRGAAP
'-"		١.,	5555	"		GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT
		l		1		FKRGLLLSALSYLGFETYQVISQAAVVHATA
				1		KVEEILEQADYLYESGETEKLYQLLTQYKESE
1		1				DAELLWRLARASRDVAQLSRTSEEEKKLLVY
l	ł	İ	į	}	ļ	EALEYAKRA/L/EKNESSFASHKWYAICLSDV
			1			GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS
	1		İ			IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP
1						*FPPYEKALG\YFHRAEQVDPNFYSKNLLLLG
ł	ŧ		ļ			KTYLKLHNKKLAAFWLMKAKDYPAHTEED
			1	•		KQIQTEAAQLLTSFSEKN
791	2141	A	6434	3	1460	IALLIVDGLAWDDOGGLALLHISPSKLIL*ODS
						SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ
						*VT\CPGQDACKE*KSTAN*GG*RE**PQVLFF
						AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ
						RLQEQRQQQSGEAEALARVYSSSISNGLSNLN
•		'				NETSGTYANGSVIDLPKSEGYYNVVSGQPSP
						DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTY\SS
ł	ł	ł				FNN/GQLAPGIT/MTEIDRIAQNIIKSHLETCQY
		İ				TMEELHQLAWQTHTYEEIKAYQSKSREALW
	ł					QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ
ļ.	ł					ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG
						KYGGMQMFKALGSDDLVNEAFDFAKNLCSL
		1				QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ
						EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA
	l					VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF
						NPDCATACK
792	2142	Α	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV
						TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL
1		1				MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT
			1			LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA
1		ł				LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH
						ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH
1	1.					GAVVNESHHDALVEDIFDKEDEDKDGFISAR
702	1		-	2001		EFTYKHDEL
793	2143	Α ,	6446	3201	152	PRLKRLVVTEEDGGARPEALGKIAPRTPAELG
		Ì				ARADQELVTALMCDLRRPAAGGMMDLAYV
	1	1				CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM
						DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH
ĺ	1 1	1		İ		EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS
						MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\
]]			WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS
1						P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P
						SGQVL\TST\ESLCRLRARVALADIAFTGGGNI
					-	VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT
	}].			DILPSLFMRCTTDLNRKDKFPAITHLKFLARD
					•	MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI
	1		.		:	FQQISPVVGDKQPTILKWRILSATNDLDRVSA
1					,	VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD
						EPAMKRPRTAGPAVHLKAMQLSWTSLALVG
L	L	L	ı <u> </u>			THE STATE OF A STOCK A STOCK A STOCK AS A ST

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	•	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j		1		peptide	Į.	/=possible nucleotide deletion, \-possible
<u></u>		<u> </u>		sequence	<u> </u>	nucleotide insertion
	ļ		1	1		IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL
1		ł		ł		LEYCMVTGYDWWDILLHVQPSMVQSLVEKL
1	ļ	1				HEEYTRQTAALQQVLSTRILAMKASLCKLSP
1	1	1				CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT
						PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF
1	1	1			İ	VLDMNTLQALQQLLQWVGDFVLYLLASLPN
		!				QPCPTSEPCPTSEPSPTSEPSPTSEPSSP*SLC\G
1		1	1			SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL
					f	KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC
1						RDEGPASEPDEALVDECCLLPSQLLIPSLDWL
						PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT
1			1			LQLDGLARAPGQPKIDHLRRLHLGACPTEEC
		[[KACTRCGCVTMLKSPNRTTAVKQWEQRWIK
						NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA
		ļ	j .	ļ .		APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT
						PRSLDHLHPEDRP
794	2144	A	6490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG
1 '	2177	^	0450	410	262	GSIEPRDLRLQ*AVITPL\TPAWVTQ
795	2145	A	6499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS
1 .55	-1-7-	1 ^	0133	333	1027	SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP
						FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP
1		l		1		DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN
						STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL
		I		1		HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS
		ſ	[DTRVRAAVGAPGLPVEPLV
796	2146	A	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA
		Į	1 1			DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR
						PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM
			'			KKMNKLQNISEELQRNISLQLMSNMNISNKIR
] [i		NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR
			.			WIWHKDSCYFLSDDVQTWQESKMACAAQN
						ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE
1						DS/YSWYESG*YNQ\PSAWVIRNAPDLNNMY
1 1			1 1	ŀ		CGYINRLYVQYYHCTYKQRMICEKMANPVQ
						LGSTYFREA
797	2147	A	6507	1	881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH
				ļ		APRAASAQLEERMRDPHPGMTLQEGDCRGS
						QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ
1				i		ESALAKLLLTCCSALRPRATQARGSSRLLVAS
]]		*	WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA
			[[ĺ	į	AIWTGAVAVLAGAAAFIYEKRGGTYWALLR
.						TLLALAAFSTAIAALKLWNEDFRYGYSYYNS
'			!·	ļ		ACRISSSDWNTPAPTQSPEEVRRLHLCTSFM
1 1	1					DMLKALFRTLQAMLLGVWILLLLASLTPLWL
	.					/SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG
			1	ļ		STHASARSQVPRSAGEAAPHSRRPPGLLPHAP
						RAASAQLEERMRDPHPGMTLQEGDCRGSQT
] [VSLTMGTADSDEMAPEAPQHTHIDVHIHQES
1				- 1		ALAKLLITCCSALRPRATQARGSSRLLVASW
	1			- 1	}	VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI
				ļ		WTGAVAVLAGAAAFTYEKRGGTYWALLRTL
]				l	}	LALAAFSTAIAALKLWNEDFRYGYSYYNSAC
[[[ĺ	ĺ	. 1	RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM
1		·		l		LKALFRTLQAMLLGVWILLLASLTPLWLYC
798	2148	A	6528	912	2287	WRMFPTKGVSP VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP
'~	-170	^	0020	712	2201	RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG
	ŀ			l		LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG
						PILLA ATMENDE ATT I ASSEDILLI MALCHERO

L000 TE	Lorom	I 14-4	Leco	The street	Fn. 32.4	1 A 1 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
SEQ ID	SEQID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ĺ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
1		l		amino acid residue of	of peptide . sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i	1		1	peptide	sequence	/=possible nucleotide deletion, \=possible
		l		sequence	ĺ	nucleotide insertion
-				sequence	ļ	EVFPEGTGLPLPHSDLPTSWCGHSLOCGSQSS
				ŀ		FPPAIHENAFIVFIASSLGHMLLTCILWRLTKK
	1					HTVSQE\DGLSLAGAPRQPRRKSRTSVLRIRV
Ì		1				MVRWELSSNGNPGRGVLGLGLGLGNKLRVV
					!	GONLGL*HCVWVVWETGE*KRWRLOMGIE*
1	1	İ	ł	ł		GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF
1			i			SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL
						GPSLPQRQGREHIVVILAAPACAPFHDR*WEP
					1	REIRPSP*ELGLRGEPTLSYPASCRVIROPIP*D
						RKSYSWKQRLFIINFISFFSALAVYFRHNMYC
		į				EAGVYTIFAILEYTVVLTNMAFHMTAWWDF
]			ľ	ļ		GNKELLITSQPEEKRF
799	2149	A	6529	1	874	FFFFQRINFIEHSGSVSLLALACDLGWCEDWS
1	ĺ	1	1	(CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV
1						DEARCKESQQEAQENLREDLCLESFAKDKIL
1			,			QIIEGSEREHEETRTKQAALDGEPLGGGQLTA
1		İ				VHLHPSKEQQGQEGGERQRGARTHHWRGW
						EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T
	İ					ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV
J	J	ļ				RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL
						GLPGDPTGPVTHHAPPVSPTGASGQERRAEP
						GAVSYAHASATK
800	2150	Α	6544	2	662	SAQRWAAVAGRWGCRLLALLLLVPGPGGAS
			İ			EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG
			1			GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF
						TASKNGTYKFCFSNEFSTFTHKTVYFDFQVG
						E\THLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV
1	1	}				GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS
801	2151	A	6556	1	1319	TPCMECIKGEGLREPONLSGSQREPOTEGSM
337		· .	****	•		DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF
						KRIFLKRMPSIRESLKERGVDMARLGPEWSQP
						MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP
						POTFKVVFDTGSSNVWVPSSKCSRLYTACVY
						HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL
						SQDIITVGGITVTQMFGEVTEMPALPFMLAEF
ł		}		ļ		DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED
						VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE
		-				GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE
						DGCLALVDTGASYISGSTSSIEKLMEALGAKE
						KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT
}						SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL
						\ALGATF\IRKFYTEFDRGNNPHGFALAR
802	2152	Α	6567	13	6147	MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL
						LAVVVLLALPVAWGQCNAPEW\LPFARPTNL
						TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS
						VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG
						IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW
						DNETPICDRIPCGLPPTITNGDFISTNRENFHY
						GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND
[İ	DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
<u>[</u>						NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA
					· :	LNKWEPELPSCSRVCQPPPDVLHAERTQRDK
]				ĺ		DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG
				ļ		DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV
			·			NLQLGAKVDFVCDEGFQLKGSSASYCVLAG
						MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR
]						CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL
]						FAKLKTQTNASDFPIGTSLKYECRPEYYGRPF
			1			I TATALLE CALLET TORPE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
+	İ		ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
}						SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG
1	ľ	l				MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI -
1		ŀ				LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS
i		ł				TNRENFHYGSVVTYRCNPGSGGRKVFELVGE
						PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
Ì	ł	i				ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP
ļ	ì	i			•	RRVKCQALNKWEPELPSCSRVCQPPPDVLHA
-	j				•	ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS
	l					MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN
1	i					GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS
	· .					YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG
		l				RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD
1	1	1				LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC
						QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP
						EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP
		1				PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG
Ì		l				HSSAECILSGNTAHWSTKPPICORIPCGLPPTI
		ł	1			ANGDFISTNRENFHYGSVVTYRCNLGSRGRK
ĺ	l		1			VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN
						KCTPPNVENGILVSDNRSLFSLNEVVEFRCOP
	ŀ					GFVMKGPRRVKCQALNKWEPELPSCSRVCQ
						PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY
						DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF
						LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL
1						KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN
1						PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR
						GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE
1						LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS
						LNYECRPGYFGKMFSISCLENLVWSSVEDNC
1						RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC
1						NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII
1						SCEPPPTISNGDFYSNNRTSFHNGTVVTYOCH
						TGPDGEQLFELVGERSIYCTSKDDQVGVWSS
1	:					PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI
1						
						IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH
						CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
1				f		KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC
						EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT
						CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS
]		•				SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL
1	'					YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS
1			'		İ	QLDHYCKEVNCSFPLFMNGISKELEMKKVYH
						YGDYVTLKCEDGYTLEGSPWSQCQADDRWD
						PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI
1						ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP
		. ,				RTLQTNEENSRVLP
803	2153	Α	6574	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL
			[ĺ	j	LLLPPLLLLLRG\SHAGNLTVAVVLPLANTSY
		1				PWSWA\RVGPAVELALAQVKARPDLLPGWT
		l	.		ļ	VRTVLGSSENALGVCSDTAAPLAAVDLKWE
		İ				HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL
						TAGAPALGFGVKDEYALTTRAGPSYAKLGDF
						VAALHRRLGWERQALMLYAYRPGDEEHCFF
]	.	LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT
	1	ł	1	l	İ	RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA
				ļ		GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW
						ERGDGQDVSARQAFQAAKIITYKDPDNPEYL
1	ļ					EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

NOT of mustication of the control of	SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence unce	, ,						
sequence Sequence 1914				l .			
uence 09/495 corresponding on the first of peptide residue of peptide residue of peptide sequence continue and peptide							
Beach Beac			l	-			
### within earl of peptide reduced peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per sequence pe			ļ				
residue of peptide sequence Poptide sequence				1			
Poptible sequence Poptible Appesible nucleotide disection Poptible			l	l			
mucleotide insertion DGLL1/1QAVTEILAHGGTV/TDGENTICRMW NRSPCQVTGYLKIDSSGORETDESL WDMDPGE NGARRAVUNNYGESL VAVSGRKLNWPLG YPPPDPRCGFDNEDPACNQDHLSTLEVLALV GLSLLGILIVSFTYTKRMQLEKELASSLWVR WEDVEPSSLERH RSAGSRLT.SGRGSNYGSL LTTEGGOFOFACTAYTKGOLLEVAKSVRKE ELLTRKVLFELKEMRDVQNEHLTRFVGACTD PPNICLT FEVCRGSLQDLENESITLDWFGRYTVAKK LUTAPELLEMASPPVRGSQAGDVYSFGILLQE LATRSGVHVEGLIDLSFEIDLPEGGHTVVAKK LUTAPELLEMASPPVRGSQAGDVYSFGILLQE ALRSGVHVEGLIDLSFEIDLPEGGHTVVAKK LUTAPELLEMASPPVRGSQAGDVYSFGILLQE GRITLRKFNRENSSINLDNLSSREGOPFR PSILALGSHLEELGILMQRCWAEDPQERPPFQ QRITLRKFNRENSSINLDNLSSREGOPFR PSILALGSHLEELGILMQRCWAEDPQERPPFQ QRITLRKFNRENSSINLDNLSSREGOPFR PSILALGSHLEELGILMQRCWAEDPQERPPFQ QRITLRKFNRENSSINLDNLSSREGOPFR PSILALGSHLEELGILMQRCWAEDPQERPPFQ QRITLRKFNRENSSINLDNLSSREGOPFR PSILALGSHLEELGILMQRCWAEDPQERPPFQ QRITLRKFNRENSSINLDNLSSREGOPFR PSILALGSHLEELGILMQRCWAEDPQERPPFQ QRITLRKFNRENSGOLBLINGHTGPVCAGV VOLKMPRYCLGEDTVATASMESSINGSPARMALAL LDAVRSFRRHRPQGCRELINGHTGPVCAGV VOLKMPRYCLGEDTVATASMESSING MERVSCLAGSTNREFERLIVEYDEGEVYKELP LVVAVLENLDSVAQODEHQVYGLELLRDINE QLTQVFGREKALRKHAEKFIFFEDSGOPKK DLOTRVESLESGOTRQLELKAKNYADQISILEE REABLKKENYAHQDHHAVHLEET KIHQLSGSDQLESTAHSKRKERPISLGIPFL ADBOLDSVAGGGSKATTPASTANSDVA TITDTTPLKEEMEGFVKYTDAFNKSEISHEH QVAQETRNVSTIGSAENEKSEVQAIBESTFEL DDMKDLSGVKSSTPTKGGENKAPDRITESL FFELSSAGSGLIGDVBGGABLLGMGREVENLI LENTQLLEFTNALINKVRDLLAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEGLRKA RABADBAGKKRSDDDBDTPAGRKRFTVVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENBAMQEKKRSSSWQFSRTSSSSNTTK KREPPYNLKYNAPTSHVTPSVKKRSSTLSQLP GOKSKAPDLSETTEALSARREQKREQVNO VKAHVQREDGRVQAFGWSLSSSNTTK KREPPYNLKYNAPTSHVTPSVKKRSSTLSQLP GOKSKAPDLSETTEALSARREQKREQVNO VKAHVQREDGRVQAFGWSLSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK KREPPYNLKYNAPTSHVTPSVKRSSSNTTK MAGRAGARDPAGRETDYPAGGE GENKRAMSHAPPAGRAGERTDYNDSS FFCNSHTRANDAGRAGERTDYNDS		[1	[
BOGLLLYIQAVTETLAHGGTVTOGENTIQRMW NSSPQGTGTYLKUSSGRETDESL WDMDPP NGAFRVVLNYNGTSQELVAVSGRKLNWPLG VPPPDPFQGFDNEPPGANODHSTLEVALV GSISLIGILVSFFYRRMQLEKELASELWRVR WEDVEPSSLERH ISAGSRLTLSGRGSNYGSL LITTEGQPQVFAXTAYYXGNL VAVKRVRKR ELTRKVLFLEKMENDQNEHLTRFYGACTD PPNICH TEYCFRGSLQDHENESTILDWAFRY SLTNDIVKGMLFLINDQNEHLTRFYGACTD PPNICH TEYCFRGSLQDHENESTILDWAFRY SLTNDIVKGMLFLINDQNEHLTRFYGACTD PPNICH TEYCFRGSLQDHENESTILDWAFRY SLTNDIVKGMLFLINDQNEHLTRFYGACTD PPNICH TEYCFRGSLQDHENESTILDWAFRY SLTNDIVKGMLFLINDQRCWAEDPGGHTVYAKK LVTAFRELLMASSPYVGSOGQDVYSFGILQE IALRSCVFHVEGLDISPKEHERVTRGEOPFFR SPLALQSHLEBLGLINDQRCWAEDPGGPFPFQ QIRLTLRKFNENSSILDNLISBMEQYANNI ELVSERCTOATLESKKABALI VALL LUAVSSFRRHIRNPGGLALEVIGHTSVAG GUKRGETVQAAFDSVTYFSDIVGFTALSAE STPMQVVTLJNDLTVTCDAVUDNEPVA VGT IGDAYMVVSGLAFDSVTYTSDIVGFTALSAE STPMQVVTLJNDLTVTCDAVUDNEPVA VGT IGDAYMVVSGLAFDSVTYTSDIVGFTALSAE STPMQVVTLJNDLTVTCDAVUDNEPVA VGT IGDAYMVVSGLAFDSVTYTSDIVGFTALSAE STPMQVVTLJNDLTVTCDAVUDNEPVA VGT IGDAYMVVSGLAFDSVTYTSDIVGFTALSAE STPMQVTVLJNDLTVTCDAVUDNEPVA VGT IGDAYMVVSGLAFDSVTYTSDIVGFTALSAE STPMQVTVLJNDLTSDAVGTSGLARVAGALIKI LLSTYTHLLGRGSTERLERGDVEMKGRG VGLKMPPYCLPFDAVTTSRAWAAL LLSTYTHLEBEGTVVTQEEPGGSGAV MSENVSGLAGSTRAFERLINDVEEVVKELIP LVVAVLENLDSVFAQDDEHQVFLELLRDDNE QUTTQ YEBERALRKAHERKFFFEDSGCORK DLOTTVESLESGUTGLELKAKNTADQISILEE REALLKEKENNALHORITBMINTWHEHENT KLHQLSGSDQLESTAHSKRKKERFISLGIFFLP AODGLITTPOAKGGFTGSGCWKFQELSOP SHTSLEDELSDVQGGSKATTTASTANSDVA TITTDTTPLKERMEGFVKVTDAVMXSEISHHE QVAQETRAVYGSALKLEEKNRELEEELRKA REARDARGKARDDDDDDTTAGRKRFTVE MARVLMERNQYKERLMELDGAVWTEMIR ASRENPAMQEKKRSSVGFTSLTSSSSNTITK KPEPPVNLKYNAPTSVKKRSSTLSQLF GDKSKAAFDT-SEETERASLASRREKKREVQVN VKAHVQKEDGRVQAFGWSLFSSNTIK KPEPPVNLKYNAPTSVKKRSSTSLSQLF GDKSKAAFDT-SEETERASLASRREKKREVQVN VKAHVQKEDGRVQAFGTLYDGSSNTIK SEPTOLSHVHLLGASTYPGDELVSSG QVDKASLCOSMTSNSSAETDSLLGGTTVVGC SAGOYTGAATSPTSTNGAREDTYDJSSND SDAYKDQISVLFREDDLYREEAQKMSSLLFT MVLGQAGCLYVYNSLDAGREDTYDJSSND SDAYKDQISVLFREDDLYREEAQKMSSLLFT MVLGQAGCLYVYNSLDAGREMTLYHATTVQHLQ LAWYGDOODVSNSILDJAGRACHTVQHLQ LAWYGDOODVSNSILDJAGRACHTVQHLQ		,					
NRSPQGTTGYLKIDSSGDRETDESLWOMDPE NGAFRVUNNTNGTSQELVAVSGRKLNWPLG YPPPDIPKCGFDMEDPACNODHLSTLEVLALV GISLIGILIVSFITYRMQLEKELASELWAVR WEDVEPSSLERHRAGGSRLTLSGRGSNYGSL LTTEGGFQVFAKTAYKGNLEVKRVNRKR ELLTRKVLFELKHMRDVQNEHLTRFVGACTD PPNCLT TEVCPGRGLODLENESTILDWFGGNYTVAKK LUTAFELLHMASPPVRGSQAGDVYSFGILLQE LARSGVHVEGLIDLSPENDLPFGGHTVYAKK LUTAFELLBHASPPVRGSQAGDVYSFGILLQE LARSGVHVEGLIDLSPENDLPFGGHTVYAKK LUTAFELLBHASPPVRGSQAGDVYSFGILLQE LARSGVHVEGLIDLSPENDLPFGGHTVAKK LUTAFELLBHASPPVRGSQAGDVYSFGILLQE LARSGVHVEGLIDLSPENDLPFGGHTVAKK LUTAFELLBHASPPVRGSQAGDVYSFGILLQE LARSGVHVEGLIDLSPENDLPSQHTVYAK LUTAFELLBHASPPVRGSQAGDVYSFGILLQE LARSGVHVEGLIDLSPENDLPSQHTVYAK LUTAFELLBHASPPVRGSQAGDVYSFGILLQE LARSGVHVEGLIDLSPENDLPSQHTVYAK QUKRGETVQAEAFDSVTYFSDIVGFTALSAE STPMQVVILDHILTCPAVIDHEDVYKVET IGDAYMVVSGLYFNNGGLARGECVARMALAL LUDAVRSFRIRHRPGCLRLRIGHTGPVAGV VGLKMPSYCLFGDTVNTASRMESNGEALKI HLSSETRAVLEFEGGFLERGDVENGKGK KVRTYWLLGGRGSSTRG ***STRAVLPEDEGGFLERGDVENGKGK*** KVRTYWLLGGRGSSTRG ***DAFFGRENGTSVARHENDRAGGERICHERDNE QLITQYEERKALRKHAEKKFIPEDSGCEKK DLOTRYESLESGTTQLELKAKNY ADOJSILEE REABLKKEYNALHQRHTEMIENYMEHLERT KLHQLSGSQLEISTANSTANSDVA TIPTDTPLKEBRGFVKVTDAFNKSEISKHIEV QVAGETRNVSTGSAEREKSESVQLISGTPLP AGDGLTPDAQKGGETTOSEGWKFQELSQPR SHSTLKDGLSSONGESKSTTPASTANSDVA TIPTDTPLKEBRGFVKVTDAFNKSEISKHIEV QVAGETRNVSTGSAEREKSESVALISGPPL AGDGLTPDAQKGGETTOSEGWKFQELSQPR AGSTLANDRAGGERSTPANSTSSEISKHIEV QVAGETRNVSTGSAEREKSESVALISGPPL QUAGETRNVSTGSAEREKSESVALISGPPL QUAGETRNVSTGSAEREKSESVALISGPPL QUAGETRNVSTGSAEREKSESVALISGPPL GGKSKAFDELSEGTSALSRREQKRRQVTNQ VALNVQRENGVAGAVFYTOWCHMIR ASRENPAMGEKKRSSINGFFSR.FSSSNTTK REPEDVALVALNGRAGGETDYAGEGERICA GGKSKAFDELSEGTSALSRREQKRRQVTNQ VALNVQREDGRVAGAVFYTOWCHMEA ASRENPAMGEKKRSSINGFFSR.FSSSSNTTK REPEDVALCHASVYGAGAVCHACHERMEA SEVDENVITLABLANGTHAUTOWCHA GGKKAFDPLSEGTERALSRREMBER GGCKSAGGRAGDDDADDITARKKFTTEVE ARAGGRAGGAGDTAVOGS GGKSKAFDELSEGTSAARRAGGKAGDTAVGC GGKSKAFDELSEGTSAARRAGGKAGDTAVGC GGKSKAFDELSEGTSAARRAGGKAGDTAVGC GGKSKAFDELSEGTSAARRAGGKAGDTAVGC SAGGVTGAATSSTNAGAGETDYAGGER GGCKSAGGRAGGTAVTOGS UTCHTTENDATATOR GGCCTATTSTANGT							
NGAFRVYLNYNGTSQELVAVSGRKLINNPUL YPPPDPIKCGFDNEDPACNOPH ISTLEVILALV GSLSLLGILVSFFYRKMQLEKELASELWRY WEDVEPSSLERH RSAGSRITLSGRSNYGSL LTTEGOGYOFAKTATYKGNLVAVKRVNKK ELTRKVLFELKHMEDVONBELTRFVGACTD PPNICLTEYCPGGSLQDLENESTILDWMFRY SLTNDIVKGMLH-IHNGALCSHGNLKSSNCVV DGRFVLKITDYGLESFEDLDFEQGHTVYAKK LVTAFPLLRMASPYRGSAGADVYSFGILLQE IALKSGVFHVEGLDLSPKSBERVTRGEOPPPF PSLALQSHLEELGLMQRCWAEDPQERPPF QURLTLRKFRENSSNILDNLLSRMEGYANNL ELVVERTOAYLEKRKAEALLYDHFSVAE QLKRGETVQABAFDSVTIYFSDIVGFTALSAE STPMQVVTLINDLYTCFDAVUNDFVXVET IGDAYMVVSGLPVRNGRHACEVARMALAL LDAVSFRRHRPDGQLRLIGHTPVCAGV VGLKMPRVCLRGDTVNTASRMESNGEALKI HJSSSTRAVLEFGOGFELERGNEKMKGK KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAFGRPPVRLPMELEDGVVYQEEPGGSCAV MSERVSGLAGSIYREFRELVRYDEEVVCRLU LUVAVAVLENLDSVEAODDHGVVELLLRDDNE QLTTQYEREKALRKHAEEKFEFFEDSQBGEKK QLTTGYEREKALRKHAEEKFEFFEDSQBGEKK DLOTREVESLESSTROLELAKANYNGELLELDDNE QLTTQYEREKALRKHAEEKFEFFEDSQBGEKK DLOTREVESLESSTROLELAKANYNGELLSCH KHLGLGSGDLBETAHSRIKKEPISTLGFPLP AGDGLLTTDAQKGGETPGSEQWKFQELSQPR SHTSLXCBLSDVSQCGSKATTPASTANSDVA TIPTDTPLKEENEGFVXVTDAPNKSESKHIEV QVAQETRAVTGASANHERSENGALSPR SHTSLXCBLSDVSQCGSKATTPASTANSDVA TIPTDTPLKEENEGFVXVTDAPNKSESKHIEV QVAQETRAVKQAKKLEEKFRFEELEELFR ARBELKKENJALHQRHTENSENGLIFPLP AGDGLLTTDAQKGGETPGSEQWKFQELSQPR SHTSLXCBLSDVSGCGSKATTPASTANSDVA TIPTDTPLKEENEGFVXVTDAPNKSESKHIEV QVAQETRAVKQAKKLUBERNRELEELFR ARBELKENSTYPKGENKACHPRITESL FFELSSAGSGLIGDVDEGADLLGMGREVFRIV MARVLMEENNQVKERLMELGEAVRAVETHER ASRENFAMGEKKRSSTNOFFSELKSSSNTITK KRPEPYNILKYNAPTSHVTSVKKRSSTLSQU VKAHVQKEDGRVQAFGWSLPQKYRQVTNG QGENKRMINLEYPYLRPIDEKDTSMLTWG VKAHVQKEDGRVQAFGWSLPQKYRQVTNG QGENKRMINLEYPYLRPIDEKDTSMLTWG GGSKATGNGGSVVGASSYFKDWAGL TGSKGRASSSSLDKLDQELKEQQKELKNQ EELSSLVWICTSHASATKETIDAVQPGNLIBS GVYCHNHULDLGRPHTSMCMTVVHDKVWGG SAEGVTOAATSSTSTNGASFVMMCRESTSSSNTIT MNEGAGNGGLVFRAGORGSVCHASSELLED BDAYKDGISVLPASSQAMSHSKESTLSKLD BDAYKDGISVLPASSQAMSHSLESTBLLYHTHOKWOG VKNNTYVQPKAMKERSSTDAHPRRCSOVG BLSYVHLLDLGRPHSISRCMTVVHDKVWGG VKNNTYVDQKAMKERSSTDAHPRRCSOVG LAWVGDVWYSHLDLGDGFHSISRCMTVHDKVWGG VKNNTYLDLGGFHISRCMTVVHDKVWGG							NRSFOGVTGYLKIDSSGDRETDFSLWDMDPF
SPPPPDPKGFENBEPACNQDHLSTLEVILAL GSLSLIGHLIVSFFYKMQLEKELSEL WRVR WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSL LITEGOFQVPAKTAYYKONULAVKRVNYRKR IELTRKVLFELKHMRDVQNBEHLTRFVGACTOR PRINCIL TEYCPRGSLQDLEWESTLDWMFRY SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV DGFFVLKITDYGLESFRDLDPEGGHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGILIQE ALLSGVFHVEGLDLSKSEHERVTRGEQPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFR PSLALQSHLEELGFLMYRGEQVANNL EELVEERTOAYLEEKRKAALLVQLDHSVAE QLKGETYQALAEDSVTYTSDIVGTTALSAE STPMQVVTLLNDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNOGHLACEVARMALAL LDAVKSFRRIERPGSQLRLINGINTDPVCAOV VGLKMPRYCLFGDTVNTASRMESNGEALLW HJSSSTRAVLEEPGGFELLERGDVENKGKG KVRTYWLLGERGSSTRG WSERNSGLAGSIYREFERLIVRYDEEVVKELD LVVAVVLENLDSVEAOPGGHELVEGVENKGKG KVRTYWLLGERGSSTRG WSERNSGLAGSIYREFERLIVRYDEEVVKELD LVVAVVLENLDSVEAOPGGHELVEGLELKRDNE QLTTQYERKALRKHAEEKFEFFEDSQGGGKAT KLHQLGSGDQLESTANSTANSTNAENVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHEEV QVAGTRNVSTIGSAENEKSESVGJGEPLP ADGGLLTTDAQKGGETTOSEQWKFQELSOPP SITISLKDELSDVGGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHEEV QVAGTRNVSTIGSAENEKSESVGJGETSTELSIENTEV QVAGTRNVSTIGSAENEKSESVGYREFTSL FELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DMDKDLSGYKGSSTPTKGIBLSTFEL DMMDLSGYKGSSTPTKGIBLSTFEL DMMDLSGYKGSSTPTKGIBLSTFEL DMMDLSGYKGSSTPTKGIBLSTFEL DMMDLSGYKGSSTPTKGIBLSTFEL DMMSLSSKAFDFISSESTNTK KYEPPVALEVNAPTSHTYBSVKKRKELESELRKA RABDARQKAKDDDDSDIPTAGRKEFTRVE MARGANGGLVYGGAFGNYPAGFDLSTSSSNTIK KYEPPVALEVNAPTSHTYBSVKKRKELESELRNQ GGENKMKNLPPYVLEPLDEKDTSMALWC QGENKMKNLPPYVLEPLDEKDTSMALWC VAHVVKEDGGVAGOSVAGSVYGDVAYDAGCL VGANLSGKTROGSVVGASVYGDVAGOL GGENKMKNLPPYVLEPLDEKDTSMALWC SABOVTGATSSTRANSAGREDIVADIS GGENKMKNLPPYVLEPLDEKDTSMALWC GGENKMKNLPPYVLEPLEGROTSMAL SEVDENYTTAEGATATATEGVDGQW DLSNYHLLDLGRPHISIGRMTVVHKVWCOG YRNKTYVVQKAMKERSTDAHPRKESOVWG LAWVGGVWYSHLDJOLGRPHISI	1						NGAFRVVLNYNGTSOELVAVSGRKLNWPLG
GSLSLLGILVSFTYRKMQLEXELASELWY WEDVEPSSLERH RSAGSKILTSERGSNYGSL LTTEGOPOVPAKTAYYKGNLVAVKRVNYKK ELTRKVLFELKHARDVQNEHLTRPVGACTD PPNICILTEYCPRGSLQDLENESTILDWAPKRVNTKK ELTRKVLFELKHARDVQNEHLTRPVGACTD PPNICILTEYCPRGSLQDLENESTILDWAPKRVSTK SLTNDIVKGMLFLHANGACSHGNLKSSNCVV DGRFVLKITDYGLESFEDLDFEQGHTVYVAK LVTAPELLRMASPVRGGQAGDVATVAKK LVTAPELLRMASPVRGGQAGDVANNI ELLVERRYGAYLEEKRKAEALLYQLIPHSVA QRIKTRKPNRENSSNLDMLLSRGVAND EELVERRYGAYLEEKRKAEALLYQLIPHSVA QLKRGETVQALAFADSVITYSDIVGTTALSAE STPMQVVTLINDLYTCFDAVIDNFDVYKSVT GIDAYMVVSGLFVRNGRHAGEVARMALAL LDAVRSFRRHRPGEQLRRIGHTGPVCAGV VGLKAPRVCLFGTVNTASRNESNGEALW HLSSETKAVLEEFGGFELERGDVEMKGKO KVRTYWLIGERGSTRG 804 2154 A 6585 2 3837 DAFGRPFVRFTRLWAYDOPEVVKELP LVVAVLENLDSVFAODDEHQVELELLRDDD QLITOYERSKALRKHAEEKTEFEDGGEGEKK DLQTRVSSLESGTRQLELKANYADOSILEE RABELKEYNALHGRITEMENPHAEHER KLHQLSGSDOLESTAHSRIKERPISLGFIPL AGGGLITPDAQKGGFTPOSSDWKFGLSOPP AGGGLITPDAQKGFTPOSSDWKFGLSOPP AGGGLITPDAQKGGFTPOSSDWKFGLSOPP AGGGLITPDAQKGFTPOSSDWKFGLSOPP AGGGLITPDAQKGFTPOSSDWKFGLSOPP AGGGLITPDAQKGFTPOSSDWKFGLSOPP SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHEPL QVAQETRNVSTGSABNEKSERVLGENSLGFPLL DMDKDLSGYKGSSTFYGLECKAKNYADOSILEE REABLKEYNALHGRITEMENPHAEHER KLHQLISGSDOLESTAHSRIKERPISLGFIPL DMDKDLSGYKGSSTFYGLECKAKNYADOSILEE RABELKEYNALHGRITEMENPKFURGLESCHK VALHENGVKROBLYVGCGFTPOSSDWKFGLSOPP AGGGLITPDAQKGFTPNSTGWKFGRISGAPP AGGGLITPDAQKGFTPNSTGWKFGRISGAPP AGGGLITPDAQKGFTPNSTGWKFGRISGAPP AGGGLITPDAQKGFTPNSTGWKFGRISGAPP AGGGLITPDAQKGFTRNSTGNAFTRVEMR ASRENPAMGEKKSSINOFFSRLESSSINTUK KPEPPVNLEVNAPTSHVTPSVKKKRSSILSSSLINGLD GGKSKAFDFISSETTASLASRRECKREFOVRO VXAHVOKEDGRVQAFGWLEGGVKGVTNG VXAHVOKEDGRVQAFGWLEGGVKGVTNG QGENSMKNLPYPVTLRPLDEKDTSMLUGGTVVGC SAEGVTOAATSPSTNGASPVMDVPPEMEAEN SEVDENNYTAEEATGAAGSAEDTVUGC SAEGVTOAATSPSTNGASPVMDVPPEMEAEN SEVDENNYTAEEATGAAGSAEDTVUGC SAEGVTOAATSPSTNGASPVMDVPPEMEAEN SEVDENNYTAEEATGAAGSAEDTVUGC SAEGVTOAATSPSTNGASPVMDVPPEMEAEN SEVDENNYTAEEATGAAGSAEDTVAGLOUNG ULSNYHLLDLGRPHISTBCMTVVHOKVOCO YNNKTYVOCPKAMKERSSTOAHPRKESGVWG ULSNYHLLDLGRPHISTBCMTVVHOKVOCO YNNKTYLDLGGVFTHAUTOHUC							
WEDVERSLENHERSAGSILTLSGRGSNYGZEL LTTEGOFOVPAKTAYYKGNU AVKRYNYRKR IELTRKVJEELKHMEDVONDEH ITRFVGACTOR PRINCILTEYCPRGES (ODIENESTILD WAMFRY SILTNDIVKGMLFLI HNGAICSHGNLKSSNCVV DGRFVLKITDYGLESFRDILDFEQGHTVYAKK LWTAFELLRMASPPYRGSQAGDVYSGIILQE IALRSGVFHVEGLDLSKEBIERVTRGEOPPFR PSILALQSHLEELGLLMKGRWAEDPQERPPFQ QIRLTLRKFNERNSSNILDNILSRMEQYANNI EELVEERTQAYLEEKRKAEALLYQULHSVAE QILKRGETVQAEADDSVTITYSDIVGFTALSAE STPMQVVTLINDLYTCFDAV UDRIPDVYKVET IGDAYMVVSGLPVRNGRIHACEVARMALAL LDAVRSFRIRHPGDQLRILIGHTVOYAGV VGLKMPRYCLFGDTVNTASRMESNGEALKI HJSSSTRAVLEEFGOFELLEROPKWGKG KVRITVWLIGERGSSTRG KVRITVWLIGERGSSTRG VARTAVILGERGSSTRG LUVAVLENLDSVFAQDQBHQVELELLRDDBV QUTYPTERKALRKHEEKTEFFEDSGEGEKK DLQTRVESEGGFELLEROPKGKG KVRITVWLIGERGSSTRG LUVAVLENLDSVFAQDQBHQVELELLRDDBV QUTYPTERKALRHEEKTEFFEDSGEGEKK DLQTRVESELSGTTQLELKAKNYADQISILEE REAELKEYNALHQRHTEMEHNYMEHLERT KLHQLSGSDQLESTAHSRIKERRISLGIFPP AGDGLITTPAQKGGETFOSSGWKYCPLSD AGDGLITPDAVGGGTFOSSGWKYCPLSD SHTSLKDELSDVSQGGKATTPASTANSDVA TIPTDTPLKEENEGGFVKVTDAPNKSISKHIBV QVAGETRNSVSTGSENGEKSEVQALESTPEL DMMKDLSGYKGSSTTYKGIERKAAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMALNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMALNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMALNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMALNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMALNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMALNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMALNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGDVDGGADLLGMGREVENLI LENTQLLETMANLNIVKRDILAKAPPNTESL FFELSSAGSGLIGD							
IBLITRKVI, FELKEMRDVQNEHLTREVQACTD PPNICLI TEVERGSLQDILENSITLD WAFREY SITNDIVKGMLFLHNGAICSHONLKSSNCVV DGREVLKITDYGLESPROLDEGHTVYAKK LWTAPELLRMASPPVRGQAGDVYSFGIILQE IALRSGVHVEGLDLSPKEIIERVTRGGPPFR PSLALQSILLEELGLLMQRGWAEDPQERPFPQ QIRLTLRKFRRENSSNILDNLLSRMEQVANNL EELVEERTQAVLBEKKKAEALLYQLPHSVAE QLKRGETVQAEAPDSVTIVFSDIVGFTALSAE STFMQVVILLDLIVTCDAVIDINFDVYKVET IGDAYMVVSGLPVENGRLHACEVARMALAL LDAVRSFRIRHRPQEQLBLRIGHTGPVCAGV VGLKMRPYCLFGDTVVTASRMESNGEALKI HJSSETKAVLBEFGGFELELGPVENGKGK KVRTYWILGERGSSTRG BOAGNEYPSTERBLEGGVVYQEEPGGSGAV MSSENSGLAGSIVREFERLIVRVDEEVVKELD LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAERFFEEDSGECKK DLQTRVESLESGTROLEKARNYADQISLEE RRAELKKEYNAHLGRHTEMHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIFPLP AGDGLTTPAQKGGETTGSEGWKFQELSOPR SHTSLKDELSDVSQGSKATTPASTANSDVA TIPTDTPLKEEMFGFVXVTDAPKSEISKHIEV QVAQETRNYSTGSAENEKSEVQAIEBSTPEL DMMKDLSGVKGSTPTKGENAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLESTKNAHVENLTEVRYEE VQVAQETRNSTGSAENEKSEVQAIEBSTPEL DMMKDLSGVKGSTPTKGENAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLESTKNAHVENLTEVRYEE MARVLMERNQYKERLMELQEAVRWTEMIR ASSENPANQEKKRSSINGFSRIFSSSNTIK KREPPVNLKVNAPISHVTTSVKKRSSTLSQLP GKSKAFDFLSEETBASLASRREQKERGVRQ VKAHVQKEDGRVAPGSHLPGKYKQVTNG QGENKMKNLPVPVYLRILDEKDTSMKLWCA VGVNJSGGKTREGGSVYGORSFYYKDVAGLD TGGSKQRSAGSSLDKLDQELKEQOKESLNO EELSLVWICTSTHSHATKVLIDAQVGFONLDS FTVCNSHVLCLASVFGARETTVPAGEDLAGGTVVGC VKAHVQKEDGRVAFTSHATTSVLIDAQVGFONLDS FTVCNSHVLCLASVFGARETDVPAGEDLSEGG QVDKASLCGSNTNSSSETBELGGTVVGC VKAAPQKEDGRVAFTSHATTSVLIDAQVGFONLDS FTVCNSHVLCLASVFGARETDVPAGEDLSEGG QVDKASLCGSNTSNSSAETBELGGTVVGC VKARCHPURPEDLAFREGORVGROBEN SEVEDENVFTLAEBETATETBOHAGSAEDTVPDIS GIFTVCNSHVLCLASVFGARETDVPAGEDLSEGG QVDKASLCGSNTSNSSAETBELGGTVVGC VAGETHNSTROASPVMDKPPENAEN SEVEDENVFTLAEBETATETBOHAGSAEDTVVDIS GIFTVCNSHVLCLASVFRAEAGAMSSLIPT MWLGAQNGCLYVHSSVAQWKCLHSIKLEC SLSVIPTTREBATTATETGOHAGSAEDTVVDIS GIFTVCNSHVLCLASVFRAEAGAMSSLIPT MWLGAQNGCLYVHSSVAQWKCLHSIKLEC SLSVIPTTREBATTATETGOHAGSAEDTVVDIS GIFTNYHLDLLGRTHHSINCENTATHTYQHEL CHARLTHATATOHLQ LAWWGDVWVYGRLASHTLATHTYQHLQ LAWWGDVWVYGRLASHLD						•	
PPNICIL TEYCFRGSLODILENSITLDWIRFRY SILTNDIVKGMILFINGAICSHGNILKSSNCVV DGRFVLKITDYGLESFRDLDFEQGHTVYAKK LWTAPELLRMASPPVRGSQDYSFGIILGE LALRSGVFHVEGIDLSPKEIIERVTRGEQPFR PSLALQSHLEELGILMQRCWAEDPQERPFPQ QIRLTLRKFRRENSSNIDNLISRMEQVANNI, EELVEERTQAYLEEKKAEALLYQU,PHSVAE QILKGETVQAEAPDSVITTSDIVGFTALSAE STFMQVVTILJNDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLYPVRNGRIHACEVARMALAL LDAVRSFRRHRPQEGLRIRGHTGPVCAGV VGLKMRYCLFGDTVTNTASMESNGEALKI, LDAVRSFRRHRPQEGLRIRGHTGPVCAGV VGLKMRYCLFGDTVTNTASMESNGEALKI, HJSSETKAVLJEFFGGFELELRGDVEMKGKG KYRTYWLLGGRGSSTRG KYTYWLLGGRGSSTRG QUTQYEEKALRKHAEKFFERILVRYDEEVVKEUP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QUTQYEEKALRKHAEKFFERILVRYDEEVVKEUP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QUTQYEEKALRKHAEKFFERISGJGFPLP AGGGLLTPDAQKGGETFGSEQWKFQELGQPR SHTSILXDELSDSVGTQGLEKARNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLGSTAHSRIKKEPPISLGIFPLP AGGGLLTPDAQKGGETGGSEQWKFQELGQPR SHTSILXDELSDSVGGGSKATFYASDIVA TIPTITPLKEEMEGFVKVTDAPPKSEISKHEV QVAQETRNVSTGSAENEEKSEVQAIGISTFIEL DMDKDLSGVKGSSTPTKGIENKAFDRIVTESL FEELSSAGGGLIGDVGRKFTRVE BELSSAGGGLIGDVGRKFTRVE MARVLMERNGYKERLMELQEAVRWTEMIR AGSRPAMOGEKKSSIWOFFSKERSOWYNG WKAHVQKEDGRVQAFGWSLPGKVKQVTING QGENKMKDLPVPVYLRPLDEKDTSMKLWGA VGWNLSGKKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASDSLDKLDQELKEQQKELNQ EELSLWWICTSTHSHATKVLIDLAVQFGMILDS FTVCNSHVLCLASVFGARETDYPAGEDLLGGTVGSND SDAYKDLSGVKTNSSAGTDSLLGGTTVYAGELSG QVDKASLCGSMTSNSSAGTBSLLGGTTVYAGELSG QUTGVTEHVFTDPLGVOOPEDLSPVYGSSND SDAYKDLSGUKTRDGGSVVGASSPND SDAYKDLSGUKTRDGGSVVGASSPND SDAYKDGGVTAASPCKORGWSLLPT MWLGAQNGCLYVHSSVAQWKCLHSIKLKD SILSIYHVKGIVLVALADGTLAFHREVOOQW DLSNYHLDLGRPHBISKCMTVVHDKVWCQ YRNKIYVVQFKAMKILESKSDAAPRKESGVVQ LAWVGDGVWYSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC							LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR
PPNICIL TEYCFRGSLODILENSITLDWIRFRY SILTNDIVKGMILFINGAICSHGNILKSSNCVV DGRFVLKITDYGLESFRDLDFEQGHTVYAKK LWTAPELLRMASPPVRGSQDYSFGIILGE LALRSGVFHVEGIDLSPKEIIERVTRGEQPFR PSLALQSHLEELGILMQRCWAEDPQERPFPQ QIRLTLRKFRRENSSNIDNLISRMEQVANNI, EELVEERTQAYLEEKKAEALLYQU,PHSVAE QILKGETVQAEAPDSVITTSDIVGFTALSAE STFMQVVTILJNDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLYPVRNGRIHACEVARMALAL LDAVRSFRRHRPQEGLRIRGHTGPVCAGV VGLKMRYCLFGDTVTNTASMESNGEALKI, LDAVRSFRRHRPQEGLRIRGHTGPVCAGV VGLKMRYCLFGDTVTNTASMESNGEALKI, HJSSETKAVLJEFFGGFELELRGDVEMKGKG KYRTYWLLGGRGSSTRG KYTYWLLGGRGSSTRG QUTQYEEKALRKHAEKFFERILVRYDEEVVKEUP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QUTQYEEKALRKHAEKFFERILVRYDEEVVKEUP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QUTQYEEKALRKHAEKFFERISGJGFPLP AGGGLLTPDAQKGGETFGSEQWKFQELGQPR SHTSILXDELSDSVGTQGLEKARNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLGSTAHSRIKKEPPISLGIFPLP AGGGLLTPDAQKGGETGGSEQWKFQELGQPR SHTSILXDELSDSVGGGSKATFYASDIVA TIPTITPLKEEMEGFVKVTDAPPKSEISKHEV QVAQETRNVSTGSAENEEKSEVQAIGISTFIEL DMDKDLSGVKGSSTPTKGIENKAFDRIVTESL FEELSSAGGGLIGDVGRKFTRVE BELSSAGGGLIGDVGRKFTRVE MARVLMERNGYKERLMELQEAVRWTEMIR AGSRPAMOGEKKSSIWOFFSKERSOWYNG WKAHVQKEDGRVQAFGWSLPGKVKQVTING QGENKMKDLPVPVYLRPLDEKDTSMKLWGA VGWNLSGKKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASDSLDKLDQELKEQQKELNQ EELSLWWICTSTHSHATKVLIDLAVQFGMILDS FTVCNSHVLCLASVFGARETDYPAGEDLLGGTVGSND SDAYKDLSGVKTNSSAGTDSLLGGTTVYAGELSG QVDKASLCGSMTSNSSAGTBSLLGGTTVYAGELSG QUTGVTEHVFTDPLGVOOPEDLSPVYGSSND SDAYKDLSGUKTRDGGSVVGASSPND SDAYKDLSGUKTRDGGSVVGASSPND SDAYKDGGVTAASPCKORGWSLLPT MWLGAQNGCLYVHSSVAQWKCLHSIKLKD SILSIYHVKGIVLVALADGTLAFHREVOOQW DLSNYHLDLGRPHBISKCMTVVHDKVWCQ YRNKIYVVQFKAMKILESKSDAAPRKESGVVQ LAWVGDGVWYSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC UAWGDGVWSRLDSTLALHTYQHDC							IELTRKVLFELKHMRDVQNEHLTRFVGACTD
DGRFVLKITDYGLESTRDLDFEQGHTVYAKK LWTAPELLRMASPPYRGSQDYSGIILGE LALRSGVPHVEGIDLSPKEIBERVTRGEQPPER PSLALQSILEELGILMQRCWAEDPOERPPFO QRRTIRKFRENSSNIDNLLSRMEQVANNI EELVEERTQATLEERKRAEALLYQUIPHSVAE QLKRGETVQAEAPDSVITTSDYGFTALSAE STPMQVVTLLNDLYTCPDAVIDNEDYXVET IGDAYMVVSGLYPUNGGILACEVARMALAL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMRYCLFGDTVNTASEMESNGEALKI H.SSETKAVLEFFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG KVRTYWLLGERGSSTRG QUTQYEEKALRKHAEKFFEEDSQCEKK DLQTTQYEEKALRKHAEKFFEEDSQCEKK DLQTTQYEEKALRKHAEKFFEEDSQCEKK DLQTTQYEEKAGNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHGJCSSGDQLESTAHSRIKSFIGJFPLP AGGGLLTPDAQKGGETFGSEQWKFQELSQPR SHTSLKDELSDVSQGSKATTDSAANSDVA TIPTDTPLKEENEGFVXVTDAPHKSEISKHEV QVAQETRNVSTGSAEPERSSTATSSDVA TIPTDTPLKEENEGFVXVTDAPHKSEISKHEV QVAQETRNVSTGSAEPERSSADGLIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGGGLIGDVDGADLLGMGREVENLI LENTQLETEINALNIVKNDLLAVEDLTCEK DVLQGELBAVRQAKLKLEEKNRELEERKA RAEAEDARQKAKADDDDSDIPTAQRECHTNI LENTQLETINALNIVKNDLAVEDLTCEK DVLQGELBAVRQAKLKLEEKNRELEERKA RAEAEDARQKAKADDDDSDIPTAQRECHTNI AGRENPAMGEKKRSSIWOFFSLESSSNTTK KFEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GGKSKAFDFI SEETEASLASREKEQVRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFVKDVAGLD TGGSKQRSASQSSLDKLDQLKEQQKELNQ EELSSLWICTSTHSHTKVLIDAVQFGNILDS FTVCNSHVLCLASVFGARETDYPAGEDLSEGG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASFVMDKPPEMEAN SEVEDENVPTAEEATATEGNAGSAEDTVDIS GTOVTTEHVFTDPLGVOFFEDLSEVGOSND SDAYKDLDSLAFFREDVATEREDQUENSEND SDAYKDLDSLAFFREDVATEREDQUENSEND DSAYKDQISVLPHEDOLJSPLVGOSND DSAYKDQISVLPHEDOLJSPLVGOSND DSAYKDQISVLPHEDOLJSPLVGOSND DSAYKDQISVRINGSLLGGTTVVGOW DLSNYHLDLGGRPHBISKGNTVHDKVWCQ VRNKIYVVQFKAMKEKSFDAHPRKESQVRQ LAWVGDGVWYSRLDSTLALTHTYQHDC	1						
LUTAPELLRMASPPYRGSQAGDVYSFGIILOGE LALRGSVPHVEGLLD.SKELIERVTRGGPPRP PSI.ALQSHLEEL.GLLMQRCWAEDPQERPPFQ QRILTLRKPNRENSNILDNLLSRRREGYANNI. EELVEERTQAYLEERKAKAALLYQULPHSVAE QLKRGETVQAEAFDSVTIYYSDIVGFTALSAE STFMQVVTLINDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNGRILHAGEVARMALAL LDAYRSFRIRIPPGGOLIRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALWI HLSSPETKAVLEFRGGFELRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALWI HLSSPETKAVLEFRGGFELRIGHTGPVEMKGKG KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVDELELRDDNE QLTTQYEERKALRKHAEKFIFFEDSQEGEK DLQTTYEERKALRKHAEKFIFFEDSQEGEK DLQTTYESKALRKHAEKFIFFEDSQEGEK DLQTTYESKALRKHAEKFIFFEDSQEGEK DLQTTYESKALRKHAEKFIFFEDSQEGEK CHQTYPTDTPLKEENEGFVKVTDAPINKSEISKHIEV QVAQETRNVSTGSAENEKSEVQAIESTFEL DMDKDLSGYKGSSTPTKGIENKAPGRNTESL FEELSSAGSGLIGDVDEGAKATGRNTESL FEELSSAGSGLIGDVDEGAKATGRNTESL FEELSSAGSGLIGDVDEGAKVAGKLELEHCH LENTQLLETKNALNVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEKKNPELEEELRKA RABAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRTEMIR ASRENPAMGEKKSSIVGFFSRLFSSSSNTIK KPEPPVNLKYNAPTSHYTPSVKKRSSTLSQLP GDKSKAFDFLSEETGALSLASREGKREGYRQ VKAHVQKEDGRVQAFGWSLPGKVKQVTNG QGENKMKNLPPYVLRIDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVPYKDVAGLD TEGSKORSASQSSLDKLJAERGQKEGVINDIS GTGYVTEHVFTDPLGWOTSMALWCA VGVNLSGGKTRDGGSVVGASVPYKDVAGLD TEGSKORSASQSSLDKLJDELKEQLSEG QVDKASLCGSMTNSSASTDSLLGGITVVGC SAEGVTGAATSPSTNGASPYMDKPPMAEAN SEVDENVPTTAECHAETATGNAGSAEDTVADIS QTGVYTEHVFTDPLGWOTSMALWCD ULSNYHLUGRPHISKERGTVVORGSVMG SLISVHVKGIVLVALADGTLAIFIKEKDVRQ DLSNYHLIDLACPHISKERGTVVORDW DLSNYHLIDLACPHISKERGTVVORDWCCL YRKKIYVVQPKAMKIEKSFDAHPRKESQVRQ DLSNYHKGIVLVALADGTLAIFIKEKDVCQ ULSNYHKGIVLVALADGTLAIFIKEKDVCQ ULSNYHKGIVLVALADGTLAIFIKEKDVCQ ULSNYHKGIVLVALADGTLAIFIKEKSQVRQ ULSNYHKGIVLVALADGTLAIFIKEKSQVRQ ULSNYHKGIVLVALADGTLAIFIKEKSQVRQ ULSNYHKGIVLVALADGTLAIFIKEKSQVRQ ULSNYHKGIVLVALADGTLAIFIKEKSGVRAC LAWVGBOGVWYSILLDSTLAIFIKHTYDHLQCHQL	į l					÷	SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV
I LALRSGVPHVEGLDLSPKEILERTTRGGOPPER PSILALOSHLEELGILLMORCWAEDPOERPPEPO QIRLTLRKFNRENSSNILDNLLSRMEQYANNI EELVERRTQAYLEERKAAEALI YQILPHSVAE QLKRGETVQAAAFDSYTIYSDIVGFTALSAE STFMQVVTLINDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNORHACEVARMALAL LDAVRSFRIRIRPGOLIRLIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNOFEALIXI HLSSETKAVLVEEFGGFELERGDVEMKGKG KVRITYWLLGERGSSTRG A 6585 2 3837 DAPGRPPVRLPTMELEDGVVVQEEPGGSGAV MSERVSGLAGSTYREFERLIVRYDEEVVKELIP LVVAVLENLDSVPAQDEGVELELLRDNE QLITQYEREKAIRKHAEEKFIFEFDSOROEKK DLQTRYESLESOTROLLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIFNYMEHLERT KLIRQLSGSDQLESTAHSRIKKEPISLGIFTLP AGDGLLTPDAQKGGETPGSGQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQGTRNVSTGSAENEEKSEVQAIESTPEL DMDKDLSGVKGSSTPTGKGENKAPATRESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLJAKVDELTCEK DVLQGELEAVKQAKKLEKERNELEEELRKA RABAEDARQKAKDDDDSDPTAGKRFTRVE MARVLMERNQVKERLIBGLGAVRFTREMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVINLKVNAPTSHUTPLDEKDTSMKLWCA VGVALSGGKTRDGGSVGASVFYKRYSTI.SQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKVKQVTING QGENKMKNLPVPYVLRPLDEKDTSMKLWCA VGVALSGGKTRDGGSVGASVFYKDVAGLD TEGSKORSASQSSLDKLDGELREQOKELKNQ EELSSLVWICTSTHSATKVLIIDAVOPGRILDS FTVCNSHVLCTASVPGARETDVAGC SAEGVTGAATSPSTINGASPVMOKPCHENS GGENKSKIPTPGDLVFAGEDLSEG QVDKASLCGSMTSNSSAETDSLLGGTVVGC SAEGVTGAATSPSTINGASPVMOKPCHESIKLKD SILSTVFIKGGVLVALADGTLAIFTRGVDQOW DLSNYHLLDLAGPTHSIRCKTVVDGCV YRNKIYVVQPRAMKIEKSFDAHFRKSQVRQ LAWVGDGUVVSIILDSTLAILYHAHTYQBLQ							DGRFVLKITDYGLESFRDLDPEQGHTVYAKK
PSI.ALQSHLEELGILMQRCWAEDPOCERPPEQ QIRLT.RRFNENSSMIDLSRMEQVANNI EELVEERTQAYLEEKRKAEALLYQILPHSVAE QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE STFMQVVTLINDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNGRIHACEVARMALAL LDAYRSFRIRIPROGOLIRIAGIHTGFVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALKI HLSSETKAYLEEFGGFELELRGDVEMKGKG KVRTYWLIGERGSSTRIG 804 2154 A 6585 2 3837 DAFGRPPVRLPTMELEDGVVYQEEPGGGGAV SEEVSGLAGSTYRFERLIVRYDEEVVKELD LVVAVLENLDSVFAODGEHQVELELLRDDNE QLITQYEREKALRKHAEKEFIEFEDSQEQEKK DLQTRYESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIENTYMEHLERT KLHQLSGSDQLESTAHSIKEPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQLSOPR SHTSIKDELSSDVSQGSKATPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNYSTGSAENEEKSEVQALESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LEENTQLIEFKNALNIVKNILAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEELRKA RAFAEDARQKAKDDDDSDIPTAQRKFTRVE MARVLMERNQYKERILMELOEAVRWTEMIR ASRENPAMGEKKRSSIWOFFSRLFSSSNITIK KPEPPVNLKYNAPTSHVTSVKKRSSTLSQLP GDKSKAFDFLSSETTEASLASRREQKEQVRQ VKAHVQKEDGGRVQAFQVTRO VKAHVQKEDGGRVQAFQVTNO QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDGELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQGRILDS FTVCNSHVLCTASVFGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVTIAEEVATATERGNAGSAEDTVDIS QTGVYTEH-YFTDPLGIVQPFELAPRYQSSND SDAYKDQISVLPNEQDL VREFAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRCLISKILKD SILSHYHKGRIVLVALADTLAFFREGVORQ BLSNYHLDLGRPHISTRCMTVVHDKVWCQ YRNKIYVVQPKAMKIEKSFDAHFRESGOVRQ LAWVGDGUVWSVISLDSTLAJHTRESGOVRQ LAWVGDGUVWSVISLDSTLAJHTHRESQOVRQ LAWVGDGUVWSVISLDSTLAJHTHRESOVRQ	1						LWTAPELLRMASPPVRGSQAGDVYSFGIILQE
QIRLTI.RKFNRENSSNI_DNLLSRNEQV_ANNI. EELVERTOQ_Y1_EEKR_AELL_YO[I_PHSVAE QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE STPMQVVTLLNOL_YTCFDAVIDNFDVYKVET IGDAYMVVSLPVRNORIHACEVARMALAL LDAVRSFRIRHRPQEQLI.RIGHTFOPV.GOV VGLKMPRYCL_FGDTVNTNSRMESNOEALLKI HLSSETKAVL_EEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTIRG 804 2154 A 6585 2 3837 DAPGRPPVRLFTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVPAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEKK DLOTRYSELSSGTRQLELKAKNYADQISILEE REAELKEYYALHORITEMIRNYMEHLERT KLHOLSGSDQLESTAHSRIKKERYMEHLIFTP AGDGLLTPDAQKGGETPGSEQWKPQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTITTIKEENEGFVKYTOTAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQALIESTPEL DMDKDLSSYKGSSTTPKGENKAPDRNTESL FELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLLAKVDELTCEK DVLQGELEAVKQAKKLEEKNEELEELRKA RABAEDARQKAKDDDDSDPTAGKRFTIRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWOFFSRLTSSSNITTK KPEPPVNLKYNAPTSHVTPSVKRSSTS.SQLP GDXSKAPDFLSSETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLEVPYYLRPJDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVYTKDVAGL TEGSKQRSASQSSLDKLDQELKEQOKELKNQ EELSSLVWICTSTHSATKVLIIDAVOPGRILDS FTVCNSHVLCLASVPGAGRENDQELEEQOKELKNQ EELSSLVWICTSTHSATKVLIIDAVOPGRILDS FTVCNSHVLCLASVPGAGRENDQELEEGOKELKNQ EELSSLVWICTSTHSATKVLIIDAVOPGRILDS FTVCNSHVLCLASVPGAGRENGANSEPMACHPYGEBLESEG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSISTINGASPVMOKPPBMEAEN SEVDENVPTAEEATEATERGNAGSAEDTVDIS QTGVYTEHVFTDPLGIVQPEDLSPVVQSSND SDAYKDOISVLPNEQDLVFREQOKELKND EELSSLVWICTSTHSATKVLIIDAVOPGRILDS FTVCNSHVLCLASVPGAGRENGANSSLIPT MWLGAQNGCLYVHSSVAQWRKCLISKILKD SILSNFWKGINLVALADGTLAIFHRGVDOQW DLSNYHLLDLGPPHISIKGMTVVHDKVWCO YRNKIYVVQPRAMKIEKSFDAHPRKESQVRQ LAWVGBOUWVSIILDSTLAILYHAHTYQBLQ LAWVGBOUWVSIILDSTLAILYHAHTYQBLQ LAWVGBOUWVSIILDSTLAILYHAHTYQBLQ LAWVGBOUWVSIILDSTLAILYHAHTYQBLQ	j .						IALRSGVFHVEGLDLSPKEHERVTRGEQPPFR
EELVEERTQAYLEKKRAEALLYQILPHSVAE QLKAGETVQAEAFDSVITYPSDIVGFTALSAE STPMQVVTLINDLYTCFDAVIDNFDVYKVET IGBAYMVVSGLYRNGRIHACEVARMALAL LDAVRSFRIRHRPGEQLIKIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALKI HLSSETKAVLEFEGGLERGDVEMKGKG KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERNSGLASSIVRFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTQVYBEKALRKHAEFIFELDSGQEKK DLQTRVESLESQTRQLELKAKNVADQISILEE REAELKKEYNALHQRHTEMINYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIPPLP AGDGLITPDAQKGGETFGSEQWKFQELSQFR SHTISLKDESLSDVSQGGSKATTTASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRIVSTGSAENTEKSEVQAIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRITESL FEELSSAGSGLIGDVDEGADLLGMGREVENIL LENTQLLETTNALINVKNDLLAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRRA RAEAEDARQKAKDDDIDPTAGRKETTVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSKLFSSSSNITK KPEPPVNLKYNAPISHSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEGRGVAGSLPFKKVTDMACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKKVACA VGVNLSGGKTRDGGSVVGASVFKDVACID FTCVCSSFVLLCLSSVFGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNYTAEBATEATEGNAGSAEDTVDIS QTGYYTEHVFTIDPLGVQFEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNYTAEBATEATEGNAGSAEDTVDIS QTGYTTEHVFTIDPLGVQFEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNYTAEBATEATEGNAGSAEDTVDIS GUGYYTEHVFTIDPLGVQFEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNYTAEBATEATEGNAGSAEDTVDIS GUGYYTEHVFTIDPLGVQFEDLAFERGVORQW DLSNYHLLDLGRPHISIRCMTVVHDKVWCQ YRNKITVVQFKAMKIKKSFDAAHRRESQVCR LAWVGDGVVVSIRLDSTRLYHAHTTYQHLQ LAWVGDGVVVSIRLDSTRLYHAHTTYQHLQ							
QLKRGETVQARAFDSVTIYFSDVGFTALSAE STPMQVVTLIADLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNGRLHACEVARMALAL LDAVRSFRIRHRPQEQLRLRIGHTGFVCAGV VGLKMPRYCLFGGTVARSMESNGEALWI HLSSETKAVLUEFGGFELELRGDVEMKGKG KVRTYWLJGERGGSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYRFERLIVRYDEVVKELIP LVVAVLEALDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REABLKKEYNALHGRHTEMIENTYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGGGLLTPDAQKGGGFTGSEQWKPQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRIVSTGSAENBEKSEVQAIESTFEL DMDKDLSGYKGSSTPTKGIENKAFDRRTESL FEELSSAGSGLIGDVDGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RABAEDARQKAKSDDDDSDIPTAGRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSINFSSNITIK KPEPPVALKVNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQVYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVYVTRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVYFXDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLESEG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGASFPSTNGASPVMDKPPEMEAEN SEVDENNPTAEEATEATEGNAGSAEDTVADIS QTGVYTEHVTDLGVQFEDLISESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNPTAEEATEATEGNAGSAEDTVADIS QTGVYTTEHVTDLGVQGPEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNPTAEEATEATEGNAGSAEDTVADIS QTGVYTTEHVTDLGVQFEDLISESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNPTAEEATEATEGNAGSAEDTVADIS QTGVYTTEHVTDLGVQFEDLISESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNPTAEEATEATEGNAGSAEDTVADIS QTGVYTTEHVTDLGVQFEDLISESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNPTAEEATEATEGNAGSAEDTVADIS QTGVYTTEHVTDLGVQFEDLISEGG QVDKASLCGSMTRRESCQVCC SAEGVTGAATSPSTNGARFRESCQVCC SAEGVTGAATSPSTNGAFRESCQVCC SAEGVTGAATSPTRAGAFRESCQVCC SAEGVTGAATSPSTNGAFRESCQVCC SAEGVTGAATSPTRAGAFRESCQVCC SAEGVTGAATSPTRAGAFRESCQVCC SAEGVTGAATSPTRAGAFRESCQVCC ANWOOLDSVAYRICDJATLRTHATTYVHDRLVWCC YR	i i						
STPMQVVTLLNDLYTCFDAVIDNFDVYKVET IGDAYMVSGLPVRIGHACEVARMALAL LDAVRSFRIRHRPQEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRNESNGEALIKI HLSSETEKAVLEEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG B04 2154 A 6585 2 3837 DAPGRPPVRLPTMELEGGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVYLELLRDDNE QLITQVYERKALRKHAREKFIEFDSQGEKK DLQTRVESLESQTRQLELKAKNYADQISLEE REAELKKEYNALHORHTMINNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLITPDAQKGGETFGSSQWKFQELSQPR SHTISLKDELSDVSQGGSKATTTASTANSDVA TIPITDTPLKEENEGFVKVTDAFNKSEISKHIEV QVAQETRINVSTGSAGNEKSEVQAIESTPEL DMDKDLSGYKGGSTPTKGEMKAFDRNTTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDIPTAGRKRFTRVE MARVLMERNGYYERLMELOFAVRWTEMIR ASRENPAMQEKKRSSIWQFSKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSWLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSWLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVLPOKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVLPOKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVPLPOKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVPLPOKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVPLPVFACHDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGAPETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGAPETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGAPREDDYPAGEDLSESG QVDKASLCGATASPSTNGAPREDDYPAGEBLSESG QVDKASLCGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGAATSPSTNGAPREDDYPAGEBLSESG QVDKASLCGAATSPSTNGAPREDDYRDAADA AWGDQGVVSTAATAATATATTYCHUC	1						
IGDAYMYVSGLPVRNGRLHACEVARMALAL LDAVRSFRRIRPOEQLALRIGIHTOPVCAGV VGLKMPRYCLPGDTVNTASRMESNGEALIKI HLSSETKAVLEEFGGFELELRGDVEMKGKG KVRITYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVELPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFEPEDSQEQEKK DLQTRVPSLESSYTRGLELKAKNYADQISILEE REAELKKEVNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLTPDAQKGGETPGSEQWKPQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNTVKNDLIAKVDELTCEK DVLQGELAVKQAKLKLEEKNRELEELRKA RAEAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKSSWOFFSRLTSSSNITIK KPEPPVNLKYNAPTSHVTPSVKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYLRRIDEKDTSMKLWCA VGVNLSGGKTROGSVVGASVFYKDVAGDD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHYLCIASVPGARETDYPAGEDLSESG QVDKASLCCSMTSNSSAETDSLLGGITVVGC SAEGYTGAATSPSTAGSPVMKKPPEMEAEN SEVDENVPTAEEAATEATEGNAGSAEDTVADIS QTGVYTEHVFTDPLGIVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQGWRCCLHSIKLKD SILSVVYKGTVLAGGTLAGTLAFTHRGVDGQW DLSNYHLLDLGRPHISIRCMTVVHDKVWCG YRNKITYVQPKAMKIEKSSTDAHPKKESQVRQ							
LDAVRSFRIRIRPOEQLRLRIGHTIOPVCAGV VGLKMPRYCLPGDTVNTASRMESNGEALKI HLSSETKAVLEEFGGFELELRGDVEMKGKG KVRITYWLLGERGSSTRG DAPGRPVRLPTMELDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEKFIEFEGSQEKK DLQTRVESLESQTRQLELKAKNYADQISSLEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLEST AHSRIKKERPISLGIFPLP AGDGLLTPDAQKGGETTGSEQWKFQELSQPR SHTSLKDELSDVSQGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNNSTGSAENEEKSEVQAILESTPEL DMDKDLSGYKGSSTPTKGIENKAFPRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELBAVQQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKFNAPTSHYTPSVKRSSTLSQLP GDKSKAFDPLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMNIPPVPYSLKPIDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQFORILDS FTVCNSHVLCLASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAFDVMDKPPEMEAEN SEVDENVPTAEEAATEATEGNAGSAEDTVADIS QTGVYTEHVFTDPLGVOPPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQMKCLHSIKLKD SLISVTVHVGGFULAADTLAFTREGVDGQW DLSNYHLLDLGRPHISIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSRLDSTLRLYHAHTYQHLQ	1			<u> </u>			
VGLKMPRYCLFGDTVNTASMESNGEALKI HLSSETKAVLEEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRO 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSTYRFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEPEDSQEQEKK DLQTRVESLESQTRIQLELKAKNYADQISILEE REAELKKEYNALHQRITEMHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERPISLGIPFLP AGDGLLTPDAQKGGETTGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTFEL DMDKDLSGYKGGSTYKGIENKAPTRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVVQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSIWQFFSRLFSSSSNITIK KPEPPVNLKYNAPTSHYTPSVKKRSSTLSQLP GDKSKAFDFLSEETENSSSNITTK KPEPPVNLKYNAPTSHYTPSVKKRSSTLSQLP GDKSKAFDFLSEETALSASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTING QGENKMKNIPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTROGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKJLIDAVQFORILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCCSMTSNSSAETDSLLGGITVVGC SAGGYTGAATSPSTAKVLIIDAVQFORILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCCSMTSNSSAETDSLLGGITVVGC SAGGYTGAATSPSTAKVLIIDAVQFORMLDS SEVDENVPTAEEAATEATGGNAGSAEDTVJOIS QTGVYTEHVFTDPLGVQPFEDLSPVYQSSND SDAYKDQISVLPNEQDLVREAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRCCLHSIKLKD SLSVIPVKGTVLVALAOGTLAFIREGVDQW DLSNYHLLDLGRPHISIRCMTVVHDKVWCG YRNKIYVVGGVALAAGTLAFIREGSVORQ LAWVGDGVWVSILDSTLRLYHAHTYQHLQ							
### HLSSETIKAVLVEFGGFELERGDVEMKGKG KVRTYWLIGERGSSTRG \$04 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDGEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELIRDDNI QLITQYEREKARKHAEEKFIEFEDSQEQEKK DLQTRVYSELSQTRQLELKAKNYADQISILEE REAELKKEYNAHQRHTEMIFNYMEHLERT KLHQLSGSDQLESTAHSRIKERPISLGIFPLP AGDGLLTPDAQKGGTGSGAVTPASTANSDVA TIPTDTPLKEENEGFVKYTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAPDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKANLNIVKINDLLAKVDELTCEK DVLQGELAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKFTTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIVQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GIKKSKAPDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPYVLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDGELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQFONILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATIGNAGSAEDTVOIS GJGVYTEHVFTDPLGVVQFFSRLSPSSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGGLYYHSSVAQWRCCLHSIKLKD SILSTVHVKGFULALAGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSRLDSTLRLYHAHTTYCHLQ							
KYRTYWLIGERGSSTRG	i 1						
804 2154 A 6585 2 3837 DAPGRPPYRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFELIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNAHQRHTEMIFNYMEHLERT KLHQLSGSDQLESTAHSRIKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIESTPEL DMMKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLLAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEEIRKA RAEAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITIK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIDAVQFONILDS FTVCNSHVLCLASVPGARETDYPAGEDLSESG QVDKASLCSBMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEELATEGNAGSAEDTVDIS QTGVYTEHVFIDPLGWOPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREAQKMSSLLPT MVLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLALADGTLAIFHRGVDGQW DLSNYHLIDLGRPHHSIRCMTVVHDKVWCQ YRNKIYVVQFKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSRLDSTLLHAHTTVQHLQ							
MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEKFIFEDSOGQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSGUESTAHSRIRKERPISLGIPPLP AGDGLLTPDAQKGGETFGSEQWKFQELSQPR SHTSKEDELSDVSQGSKATTPASTANSDVA TIPTDTPLKEREGFVKVTDAFNKSEISKHIEV QVAQSTRINVSTGSAENEEKSEVQALIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRINTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKNDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKSSIWQFSRLFSSSNITTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEEDRVAFWSUKPKRLSVCYTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGYNLSGGKTRDGGSVVGASVYYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQVKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCLASVPGAETDYPAGEDLESSEG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEAATGANGSAEDTVVDIS QTGYTTEHVFTDPLGIVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSEVAQWKKCHSIKLKD SILSIVHVKGIVLVALADGTLAFHREVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	804	2154	Δ	6585	7	3927	
LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKEYNALHQRHTEMIENYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPAQKGGETFGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAFNKSEISKHIEV QVAQETRINVSTGSAENEEKSEVQAHIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDBGADLLGMGGREVENLI LENTQLLETKALNIVKNDLLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAAEADARQKAKDDDDSIDTAQRRKFTTEVE MARVLMERNGYKERIMELOEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQVRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYLIPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASYPYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEAATEATEGNAGSAEDTVODIS QTGVYTEHVFTDPLGVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAFHRGVDGQW DLSNYHLLDLGRPHHSIRGMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	004	2134	^	رەدە	2	3837	
QLITQYEREKALRKHAEEKFEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHORHITEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKPQELSQPR SHTSLKDELSVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RABAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRIDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEKATEATBGNAGSAEDTVJDIS QTGVYTEHVFTDPLGIVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCYYMSLAQWKKCLSILLPT MWLGAQNGCYVHSSVAQWKKCLSISLLFT MWLGAQNGCYYHSRAMGKRISILKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDCVVSIRLDSTLRLYHAHTYQHLQ	İ						
DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSQQLESTAHSRIKERPISLGIPPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGGIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEADARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTTHASTKV-LIDAVQPGNILDS FTVCNSHV-LCASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPYMDKPPEMEAEN SEVDENVPTAEEATEA-TEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDVVSIRLDSTLRLYHAHTYQHLQ	1						·
REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPOSEQWKPQELSQPR SHTSLKDELSDVSQGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNYSTGSAENEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWGFFSRLFSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVVDIS QTGYYTEHVFTDPLGVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDQQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWYSIRLDSTLRLYHAHTYQHLQ	1						
KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETTGSEQWKFQELSQFR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAFNKSEISKHEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYVLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCLASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGYYTEHVFTIDLGVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHISIRKMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	1		l				
SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPITDTPLKEENEGFVKVTDAPNKSEISKHEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTFTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA REAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIDAVQPGNILDS FTVCNSHVLCIASVFGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGIVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKTYVVQPKAMKEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	Į						
TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRINVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEADARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQPEDLSPVYQSSND SDAKKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGTVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQFKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							AGDGLLTPDAQKGGETPGSEQWKFQELSQPR
QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RABAEDARQKAKDDDSDIPTAQRKFFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVYGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIPHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							SHTSLKDELSDVSQGGSKATTPASTANSDVA
DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASSPVMDKPPEMEAEN SEVDENVPTAEEATEGNAGSAEDTVDIS QTGYYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	}	1			ļ		TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV
FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAK VDELTCEK DVLQGELEAVKQAKLKLEEKNELEELRKA RAEAEDARQKAKDDDDSDIPTAQRKFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKTYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	1						
LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVNDIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHYKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	1 1			1			1
DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLDDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	1	İ		ŀ			1
RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITIK KPEPPVILKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ]	j		j			
MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	1	[
ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKV.IIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ			Ì				
KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ]		i	j			
GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ		[[ĺ	I		KDEDDAM KAN VDLETALDSTANDSSALSON ILK
VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ		l			İ		CDK SK V EDEL SEELE V SI V SD DEUKDEUMD
QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ]	ļ		J	j	
VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ			l		Į.		
TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ]		1		
EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ]	ļ	ļ	j	j	J	
FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ		ļ	l	- 1			
QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ			l				
SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ		ļ	ŀ	ļ	. 1	1	
SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ			ļ				. -
SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ			- 1		ļ	1	1
MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ			j		į	ļ	
SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					ļ	l	
DLSNYHLLDLGRPHHSIRCMTVVHDKVWČG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ		J	i	İ	l	Ţ	
YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ		1	ł			ì	
LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ		1	ĺ		1	j	DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG
		- 1	ĺ	-	[j	YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ
DVDIEPYVSKMLGTGKLGFSFVRITALMVSC				1		l	
							DVDIEPYVSKMLGTGKLGFSFVRITALMVSC

SEQ ID NO: of No: of No: of nucleotide peptide entitle sequence (A=Alanine C=Cyste degration and peptide sequence (A=Alanine C=Cyste nucleotide location corresponding and possible nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanine C=Cyste nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanine C=Cyste nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanine C=Cyste nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanine C=Cyste nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanine C=Cyste nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Alanine C=Cyste nucleotide location (A=Amino acid sequence (A=Amino acid	LGLR VTPGT VPGQV RRP ATFTQ LVGEP IWYKS PTLLQ IDTGL FLLPT LIREV /TAPL LTCRA
nucleotide sequence	LGLR VTPGT VPGQV RRP ATFTQ LVGEP INTEL IDTGL IDTGL ITCRA IRVWE
eotide sequence USSN 09/496 corresponding to last amino acid residue of peptide sequence Poptide sequence USSN 09/496 corresponding to last amino acid residue of peptide sequence Poptide sequence USSN 09/496 corresponding to last amino acid residue of peptide sequence Poptide sequence Poptide sequence I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide mesertion NRL WVGTGNGVIISIPLTETVILHQGRI ANKTSGVPGNRPGSVIRYYGDENSDK FIPYCSMAHAQLCFHGHRDAVKFFVA ISPQSSSSGTDLTGDKGRGHLHRSLVV VRIKCALFYGYIRTNYSLAQSAGISLM SGPGDFEEPLAFDGSRMSKEEDSIWFRI DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN FOYSGOVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	LGLR VTPGT VPGQV RRP ATFTQ LVGEP INTEL IDTGL IDTGL ITCRA IRVWE
sequence Sequence	LGLR VTPGT VPGQV RRP ATFTQ LVGEP INYKS PTLLQ IDTGL FLLPT LIREV /TAPL LTCRA
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 914 15- Possible nucleotide deletion, Possible nucleotide insertion 15- NRL WVGTGNGVIISIPLTETVILHQGRI ANKTSGVPGNRPGSVIRVYGDENSDK FIPYCSMAHAQLCFHGHRDAVKFFVA ISPQSSSSGTDLTGDKGRGHLHRSLVV SLKVVTKRGSADGCTDWSIDIKKYQV VRIKCALFYGYIRTNYSLAQSAGLSLM SGPGDFEEPIAFDGSRMSKEEDSIWFRI DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKDTL REDDIGNYTCELKYGGFVVRRTTELT TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVRR	LGLR VTPGT VPGQV RRP ATFTQ LVGEP INYKS PTLLQ IDTGL FLLPT LIREV /TAPL LTCRA
amino acid residue of peptide sequence	VTPGT VPGQV RRP ATFTQ LVGEP RWYKS PTLLQ DTGL FLLPT LIREV /TAPL LTCRA
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NRLWVGTGNGVIISIPLTETVILHQGRI	VTPGT VPGQV RRP ATFTQ LVGEP RWYKS PTLLQ DTGL FLLPT LIREV /TAPL LTCRA
sequence Sequence nucleotide insertion	VTPGT VPGQV RRP ATFTQ LVGEP RWYKS PTLLQ DTGL FLLPT LIREV /TAPL LTCRA
sequence Sequence nucleotide insertion	VTPGT VPGQV RRP ATFTQ LVGEP RWYKS PTLLQ DTGL FLLPT LIREV /TAPL LTCRA
ANKTSGVPGNRPGSVIRVYGDENSDK FIPYCSMAHAQLCFHGHRDAVKFFVA ISPQSSSSGTDLTGDKGRGHLHRSLVV 805 2155 A 6605 469 2602 FGRLLWGTAFKSWKMKAPIPHLILLY, SLKVVTKRGSADGCTDWSIDIKKYQV VRIKCALFYGYIRTNYSLAQSAGLSIM SGPGDFEEPIAFDGSRMSKEEDSIWFRI DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEGVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	VTPGT VPGQV RRP ATFTQ LVGEP RWYKS PTLLQ DTGL FLLPT LIREV /TAPL LTCRA
FIPYCSMAHAQLCFHGHRDAVKFFVA ISPQSSSSGTDLTGDKGRGHLHRSLVV 805 2155 A 6605 469 2602 FGRLLWGTAFKSWKMKAPIPHLILLY SLKVVTKRGSADGCTDWSIDIKKYQV VRIKCALFYGYIRTNYSLAQSAGLSLM SGPGDFEEPIAFDGSRMSKEEDSIWFRI DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEGEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVRR	VPGQV RRP ATFTQ LVGEP WYKS PTLLQ IDTGL FLLPT LIREV /TAPL LTCRA IRVWE
ISPQSSSSGTDLTGDKGRGHLHRSLVV 805 2155 A 6605 469 2602 FGRLLWGTAFKSWKMKAPIPHLILLY, SLKVVTKRGSADGCTDWSIDIKKYQV VRIKCALFYGYIRTNYSLAQSAGLSLM SGPGDFEEPIAFDGSRMSKEEDSIWFIL DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTV LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVRR	RRP ATFTQ LVGEP IWYKS PTLLQ IDTGL FLLPT LIREV /TAPL LTCRA IRVWE
805 2155 A 6605 469 2602 FGRLLWGTAFKSWKMKAPIPHLILLY, SLKVVTKRGSADGCTDWSIDIKKYQV VRIKCALFYGYIRTNYSLAQSAGLSLM SGPGDFEEPIAFDGSRMSKEEDSIWFRI DSGLYACVIRNSTYCMKVSISLTVGEM CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEM SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIDRD. YIEDVARCVDQSKRLIIVMTPNYVRR	ATFTQ LVGEP WYKS PTLLQ IDTGL FLLPT LIREV /TAPL LTCRA IRVWE
SLKVVTKRGSADGCTDWSIDIKKYQV VRIKCALFYGYIRTNYSLAQSAGLSLM SGPGDFEEPIAFDGSRMSKEEDSIWFRI DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTTYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD	LVGEP WYKS PTLLQ IDTGL FLLPT LIREV /TAPL LTCRA IRVWE
VRIKCALFYGYIRTNYSLAQSAGLSLM SGPGDFEEPIAFDGSRMSKEEDSIWFRI DSGL YACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD	WYKS PTLLQ PTLLQ PTLLPT LIREV PTAPL LTCRA PRVWE
SGPGDFEEPIAFDGSRMSKEEDSIWFRI DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD	TTLLQ TDTGL FLLPT LIREV TAPL LTCRA TRVWE
DSGLYACVIRNSTYCMKVSISLTVGEN CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD	DTGL FLLPT LIREV /TAPL LTCRA IRVWE
CYNSKMKYFEKAELSKSKEISCRDIED REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD	FLLPT LIREV /TAPL LTCRA IRVWE
REPEILWYKECRTKTWRPSIVFKRDTL REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEGEVSISLIVDSVEEGGI CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	LIREV /TAPL LTCRA /RVWE
REDDIGNYTCELKYGGFVVRRTTELTY TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD	/TAPL LTCRA /RVWE
TDKPPKLLYPMESKLTIQETQLGDSAN FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD	LTCRA RVWE
FFGYSGDVSPLIYWMKGEKFIEDLDEN SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	RVWE
SDIKILKEHLGEQEVSISLIVDSVEEGD CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	
CYVENGNGRRHASVLLHKRELMYTVI LGAILLLLVCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	
LGAILLLUCLVTIYKCYKIEIMLFYRN EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	
EELDGDNKDYDAYLSYTKVDPDQWN EERFALEILPDMLEKHYGYKLFIPDRD YIEDVARCVDQSKRLIIVMTPNYVVR	
EERFALEILPDMLEKHYGYKLFIPDRD: YIEDVARCVDQSKRLIIVMTPNYVVRR	
YIEDVARCVDQSKRLIIVMTPNYVVRR	
[
VEALKHTIKLLTVIKWHGPKCNKLNSI	` '
LQYEMPFKRIEPITHEQALDVSEQGPFG	
VSAISMAAATSTALATAHPDLRSTFHN	• •
QMRQKHYYRSYEYDVPPTGTLPLTSIC	
YCNIPMTLINGQRPQTKSSREQNPDEA	
ILPLLPRETSISSVIW	
806 2156 A 6614 3 1584 NSARGGVGVRGARAMATVQEKAAAI	NLSAL
HSPAHRPPGFSVAQKPFGATYVWSSIN	√TLQT
	/IFSHL
IQNKYFGDVDIPRAKVVRVCQALMDY	KVFE
AVPTKVFGKDKKPTFEDSSCSLYRFTT	IPNQD
SQLGKENKLYSPARYADALFKSSDIRS	
LWENLSLKPANSPHVNISTTLSPQVINE	•
ETIGRLLQLVDLPLLDSLLKQQEAVPK	· .
RQSTMVNSSNYLDRGILKAYSDSQEDI	
AIDCLEYLPDQMVVEISRSFPEQPDRTI	
LLFDAIGRYYSSREPLLNHLSDVHNGIA	
NGKTEIALEATQLLLKLLDFQNREEFR MAVAANPSEFKLQKESDNRMVVKRIF	
DNKNLSKGKTDLLVLFLMDHQKDVF	
L\HKIVS\VK\LMAIQNGRDPNRDAGYI	
DORDYSNITEKTTIDELLYLLKTLDEDS	
KEKKKLLGQFYKCHPDIFIEHFGD	ALLON
807 2157 A 6615 4198 2094 FGIVGTFALETDELDSDRDPAIFSLCDF	GAMD
PQILLLALLTLGLAAQHQDKVPCKM/	
CPDRVDKKVSCQVLGLLQVPSVLPPD	
LSGNOLRSILASPLGFYTALRHLDLSTN	
QPGAFQALTHLEHLSLAHNRLAMATA	
GLGPLPRVTSLDLSGNSLYSGLLERLLG	
LHTLSLAENSLTRLTRHTFRDMPALEQ	
NVLMDIEDGAFEGLPRLTHLNLSRNSL	
FSLQQLRVLDLSCNSIEAFQTAS\QPQA	
WLDLRENKLLHFPDLAALPRLIYLNLS	
LPTGPPQDSKGIHAPSEGWSALFLS\AP:	SGNAS
GRPLSQLLNLDLSYNEIELIPDSFLEHLT	
NLSRNCLRTFEARRLGSLPCLMLLDLS	
TLELGARALG\SLRTLLLQGNALRDLPF	YTFA
NLASLQRLNLQGNRVSPCGGPDEPGP\\	SGCV\
AFSGITSLRSLSLVDNEIELLRAGAFLH	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	l	ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		!	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
		<u> </u>	 	sequence	<u> </u>	LDLSSNPGLEVATGALGGLEASLEVLALQGN
ĺ	1		[GLMVLQVDLPCFICLKRLNLAENRLSHLPAW
				•		TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR
			İ	ı		RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA
1	Į.	}	ļ.			TODLICRESSOEEVSLSHVRPEDCEKGGLKNI
		1	!		Ì	NLIILTFILVSAILLTTLAACCCVRROKFNOO
1]	j	YKA
808	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG
					1	ARHRNNTEKKHPGGGESDASPEAGSGGGGV
]	ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN
		1]		}	AGSVGLALIVWIVTGFITVVGALCYAELGVNI
					İ	PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP
1						TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA
			}			AICLLLLTWVNCSSVRWATRVQDIFTAGKLL
1	}					ALALIIIMGIVQICKGEYFWLEPKNAFENFQEP
	1		1			DIGLVALAFLQGSFAYGGWNFLNY\VTEELV
			}		j	DP\YKNL\PRAIFISIP\LVTFVYVFANV/ALYVT
						AMSPQEL\LAS\NAVAVTFGEKLLGVMAWIM
						PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP
						SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD
'			į ,]	MYTLINYVGFINYLFYGVTVAGQIVLRWKKP
						DIPRPIKINLLFPIIYLLFWAFLLVFSLWSEPVV CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI
						ELLTLVSQKMCVVVYPEVERGSGTEEANED
						MEEQQQPMYQPTPTKDKDVAGQPQP
809	2159	A	6621	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH
						TAGAKRYKYLRRLFRFROMDFEFAAWOMLY
						LFTSPQRVYRNFHYRKQTKDQWARDDPAFL
l l						VLLSTWLCVSTIGFGFVLDMGFFETIKLLLWV
						VLIDCVGVGLLIATLMWFISNKYLVKRQSRD
						YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN
						HVILTDTFIGYLVGNTLWLVAVGYYTYVTFL
						GYSVGLLFFS\ALPFLKNTVILLYPFAPLILLYG
						LSLALGWNFTHTLCSFYKYRVK
810	2160	A	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA
						QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP
						EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK
						PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ
						IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT
						QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF
						IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN F
811	2161	Ā	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG
"	2101	^	5027	.0	3301	AAASISQEPSLADIKARIQDSDEPDLSQNSITG
						EHSQLLDDGHKKARNAYLNNSNYEEGDEYF
	i.	l		ŀ		DKNLALFEEEMDTRPKVSSLLNRMANYTNLT
	, i			l		OGAKEHEEAENITEGKKKPTKTPOMGTFMG
				1		VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF
				l		AIVLICCCCTMLTAISMSAIATNGVVPAGGSY
	ļ	ľ	1	!	İ	FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL
				İ		GAIEIFLVYIVPRAAIFHSDDALKESAAMLNN
						MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL
				J		ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT
	[[ſ		. [LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS
		1		j		QFFNATCDEYFVHNNVTSIQGIPGLASGITEN
	ļ	i	ł		ł	LWSNYLPKGEIEKPSAKSSDVLGSLNHEYVL
						VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD
	j	į	l	Į.	I	
						AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Ī			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł				peptide		/-possible nucleotide deletion, \-possible
<u></u>	ļ			sequence		nucleotide insertion
						FGHSKANGEPTWALLLTAAIAELGILIASLDL
1	ļ	1				VAPILSMFFLMCYLFVNLACALQTLLRTPNW
	[ĺ	[f	RPRFRYYHWALSFMGMSICLALMFISSWYYA
			1		!	IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS
	i '	ŀ	1			LSAARFALLRLEEGPPHTKNWRPQLLVLLKL
						DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV
			1		•	GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ
						LVVAAKLREGISHLIQSCGLGGMKHNTVVM
						GWPNGWRQSEDARAWKTFIGTVRVTTAAHL ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG
}		ļ				GMLMLLPFLLK\QHKVWRKCSIRFF\TVAQLE
1						DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS
						DISAYTYERTLMMEQRSQMLRHMRLSKTER
1				ľ		DREAQLVKDRNSMLRLTSIGSDEDEETETYO
			[EKVHMTWTKDKYMASRGQKAKSMEGFQDL
ŀ						LNMRPDOSNVRRMHTAVKLNEVIVNKSHEA
						KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL
			[ERVLLVRGGGSEVITTYS
812	2162	A	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG
1			1 1			DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP
1						MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ
1						VAGADFESEDEGEEFDDWEDDYDYPEEEQLS
1						GAGYRVSAALEEADKMFLRTREPALDGGFQ
						MHYEKTPFDQLAFIEELF\SLMVVNRLTEELG
<u></u>						CDEIIDRE
813	2163	Α	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC
						WQYRQLSALHRAPRPTRPDKARRLGYKAKQ
						GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG
						ATYGKPVHHGVNQLKFARSLQSVAEERAGR
						HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK
						AIRRNPDTQWITKPVHKHREMRGLTSAGRKS
						RGLGKGHKFHHTIGGSRRAAWRRRNTLQLH RYR
814	2164	A	6635	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR
"	2107	^	2022	~01	1103	DSEDRSDSRAAOPAHDSGHGDDESPSTSSGT
	Ì					AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN
]					!	PLTKESIRQKEMESKRLRLLQEEDRRKKIARM
[ſ			ĺ		GFNASSMLRKSQLGFLNVTNYCHLAHELRLS
1						CMERKKVQIRSMDPSALASDRFNLILADTNS
]	1	-		l		DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF
				ļ		MHENLYFTNRKVINSVCWASLNHLDSHILLC
		Ì			l	LMGLAETPGCATLLPASLFVNSHPAGIDRPG\
				1	l	MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR
	1	ł	·	ì		RVLLTNVVTGHRQSFGTNSDVLAQQFALMA
			. [ļ	İ	PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF
	})	J			HDSAVTSVRILQDEQYLMASDMAGKIKLWD
	İ			ſ		LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL
			ļ			VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA
[[ļ		. 1			DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY
		لـــــا				SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG
	ĺ					PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE
	ł	ł	ł	i	İ	RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT
			Ì			PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF
[]			ĺ			QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA
			j			LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS
[[- 1	1	ľ	ľ	NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR
	l		ļ			VSVDWGKCMNPFRNMVLEILDVSGNGWTV
	İ		ĺ	1		DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN
L						IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

SSQ ID Stop of hod peptide coulde sequence (A-Alamine Co-Cysteine, peptide coulde sequence (A-Alamine Co-Cysteine, peptide coulde sequence (A-Alamine Co-Cysteine, peptide coulde sequence (A-Alamine Co-Cysteine, peptide sequence) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A-Alamine Color) sequence (A							
nucleo died sog- sog	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Docation Docation	1		hod	1.			
uence Month			i				
Beach Beac			i	1		corresponding	
amino acid recibited sequence periode sequence periode sequence periode sequence provide sequence provide sequence provide sequence periode se		uence	ł	1			
residue of pspidid sequence V=Tyrosine, X=Ulatowwa, **Sign ordor, **possible nucleotide floited, **possible nucleotide fl	uence	1		914			
#Possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion nucleotide insertion nucleotide insertion in publicate insertion in Professional Professiona		[1			
Sequence nucleotide insertion		1		1	1	sequence	1
RIVETIADIKVINIAYNKINKIADEAPYGID NQVINISYNLIGELYSSNYYGLKVAYIDL QKNHIAJIQODTREFLEKIGTILDIRDNALITH FPSIPDIELGSNIKLYTLEKINTANLIHLESEN LENLDILYFILKVPHIQUILINQNRESSCSGOQ TYSENSI EQLIFIGEMIQ OLAWFEILCWDVF EGLSHLQVLYJAHNYINSI,PEVPSHITALR GISINSKRITVLSINDIPANIELDISRNQIL APPIDVFVSLSVLDITHREFICECELSTFINWI, NHINVTIAGPPADIYCVYPDISLGSVLSILSISTE GCDEEVLKSLKFSLFVCTVTI.TI.FIMTILIY TERGFCFICYXTAQRIVEKDHOGTEPDMY KYDAYLCFSKDFTWVQNALLKHI.DTQYSD QNEPALCFEERDFVGDIRDNANIORALWSIR KIVCLVSRIFIRDGWCLEAFSYAQGRCLSDI. NSALIMVVOSSLSQVJCMSHQSIRGFVQKQQ YLRWPEDLQDVGWTLHKI.SQQILKKEKEKK KDNNIPLQTVATIS SALIVATORQATRETAPATMSTTILVAT ANHATRATNOQYYGCAGGAAV.COMITIT BERGLYSSQCOVOTVARHIVMFELMWEPHNY STEDLDRQNWISMFSERAAVEFINKQVCLA KONSTISLDAAAGQUFOTYARHVEPHNYEPHNY STEDLDRQNWISMFSERAAVEFINKQVCLA KONSTISLDAAAGQUFOTYARHVEPHNYEPHNY STEDLDRQNWISMFSERAAVEFINKQVCLA KONSTISLDAAGQUFOTYARHVEPHNYEPHNY STEDLDRQNWISMFSERAAVEFINKQVCLA KONSTISLDAAGQUFOTYARHVEPHNYEPHNY STEDLYNTQAGGQVVTVARHVYRELMWEPHNY VSPTSIFFKSGEPLATKSNSI.SEQLANTER AVARALISMMARMGQPBLPILPPQLDSNDSEI EDVNTLQGGQQVVTVSQARGAAVTSQLQ FYRRYCLA KOSSDCHSVSSRDSAAPSPROADNI.SADPY VSPTSIFFKSGEPLATKSNSI.SEQLANTER AVARALISMMARMGQPMLPILPPQLDSNDSEI EDVNTLQGGQQVVTVSQARGAVTSQLQ FYRRYLTAGGGGGAATAYROAAVTSQLQ FYRRITAGGGGAATAYROAAVTSQLQ FYRRITARAVSKVADKMHIMTEKELQKH SAGNSMLPSNSVTMETSKAMSNIQRIQENER LIQELLEEKSNEERONDISIELERNQKYVSQS NLMMERNNSVTMETSKAMSNIQRIQENER LIQELLEEKSNEERONDISIELERNQKYVSGS NLMMERNNSVTMETSKAMSNIQRIQENER LIQELLEEKSNEERONDISIELERNQKYVSGS NLMMERNSVTMETSKAMSNIQRIQENER LIQELLEEKSNEERONDISIELERNQKYVSGS NLMMERNSVTMETSKAMSNIQRIQENER ESQAQSKYRSKQKRRQLELVTSLEELITDL RVEKSLENNIESRIKKSAQRSQAEEIDEI RKYTQEELDKARQLIKKTNSTDQAAAGQLS LIQAELGERNBEERONDISIELERNQCHAACHAL LICHMYTIKMYTUQLLNQQCKEESSSEEEEE EKAEERRRSSQCOSAASSGQOQAAAGQLS LIQAELGERNSEERONDISTELERNULOP TSIPPEPLGPVSMSSCEESLAASPMAAKUPDN PGRWVCVREAVPOPLQESSTISLSTSUPEE GDPLALGERSDEGEDEDEVSMKGR PEPTLFGDDDDDDDDDGUGSTISLSTSUPEE GDPLALGERDDDDDDDDDUG		ŀ	1				
NILOVILNISYNLLGELYSSNYTOLERWAYDDL (NYHLARIO)DOTREFLEKLOTILDLROWALTH FPSIPDIFLSGNKLVTLEKINLTANLHLSENG LENLDLLYFLLRVPHLQULLNOWESSCSGOQ TYSENYSLEQLFLGENMUQLAWETELCWDVF EGLSHLOVLYLINNYNLSNEPOVESHLTALR GLSINSRRLTVLSINDLPANLELDISRNOLL APPEDVYSLSVLDTHREFICECELSTFINVL MHTNYTLAGPADIYCVYPDSLSGVSLSSISTE GCDEEPLYLSLKPSLPVCTVTLTLTLTLTLTV TERRGFCFICVXTAQRLYRXDHPQGTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTDYSD QREPPLCFEERDPYVGENRPANIQDAIWNSR KIVCLYSRHTLRDGWCLEAFSYAQGRCLSDL NSALBWVGSLSGYQLMENIGSIRGFVQKOQ YLRWPEDLQDVGWFLRKLSQQLKKEKEKK KNNNIPLQTVATIS BRAGVERGKQHAAAAAFYDVGGDRPWDS GRYQLPRIPRYNKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHOREFFQVTAPKQP KKQQTTAATGNQATKTAPATMSTTLTVAT AVHAYRYTNGQYVKQCKFGAAVLGNHTTR EYRLLYISQQQVTVARHINVFELWRYNNY STYDDDQROWSIMPESEKAAVEFNKQVCLA KCNSTSSLDAVLSQDLYADGPAVEYGOSLE VAYTGWLPQNHVLGQVFDSTANKDKLLRLK LOSGKYKGWEDGMOMKKGGKKLLIVPA CAVGSEGVIGWTQATDSILVFEVEVRIVKLIA LOSGKYKGWEDGMOMKKGGKRLLIVPA CAVGSEGVIGWTQATDSILVFEVEVRIVKLIA KDSGSDGHYSSRDSAAPSPPGADNISADPV VSPTTSIFFKSGEPALRTKSNIS.SEQLAINTSSD AVKAKLISMAKMGQMPLIPPGJOSDNSSEL EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTQAPPSVTGLQAFAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPOASAVTSGLQ PVRLYPALSQPPHGGSGMMSHIMEAR QNNTERMAVSKVADKMDHLMTKVEELQKH SAGNMBLPSMSVTMETSMISHNIGQENER LKQELEKSNIBEGNDKISELIERNQRYVCOS NIMMERKNISLGTAFANDLAKSLLGPT SEQAQSKFKSEKQNRKOLELKYTSLEELTDL RVEKSELERNLSEKKKKSAQERSCARALAS DPSEKVKLIMNQVVGRUKKSAGKSGVCAALAS DPSEKVKLIMNQVVGRUKKSAGKSGVCAALAS DPSEKVKLIMNQVVGRUKLOLKKTRVSTDQAAAGLOJ LVABALDTOWARKCHKKSAGLERSCARASED RKSYOEELDKLRQLLKKTRVSTDQAAAGLOJ VFRLYPARLSQPHERELEGNINGQUSCVEAAAS DPSEKVKLIMNQVVGRUKLEGNINGGNINGCHER LKQELEKSNIBEGNINGGLIPPLACTISHINGLICQHER RKSYOEELDKLRQLLKKTRVSTDQAAAGLOJ PVRLYPARLSQPHERELEGNINGGNINGCHER LGELEKSNIBEGNINGGNINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKSRINGCHER LGELAKS	<u> </u>	.	<u> </u>	<u> </u>	sequence	L	1
QKNHANQOPTKFLEKUTILDLRDNALITIH FPSIPDELGONKLYTIKRINTANLHLSENR LENLDILYFLIKVPHLQILLINQNRFSSCSGOP TYSENSPLEQLIFIGEMUQLAWFEICLCWDVF EGLSHLQVLYINHNYLNSLPPOVPSHLTALR GSLINSNRLTVLSINDLPANLELDISRNQLL APPIDVFVSLSVLDITHREFICECELSTFINWL MHTNYTLGPPADIYCVYPDLSGOVSLPSLSTST GCDEEPVLKSLKFSLPVCTVYT.IT.FLMTILTV TKFRGFCFICYXTAQRLVFXDHPQGTEPDMY KYDAYLCSSKDFTW-QNALLKALDTOYSD QNRFNLCFERDFYPGENRFANIQDAIWNSR KYCLVSRIFIRDOWCLEAFSYAQRCLSDL NSALMYVVOSLSGYQLMKHQSIKGFVQKQQ YLRWPEDLQDVGWFPKIKSQQLKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRFAGKGHAAAAFYDVGGDRFWDS GNTQLPFRIPYVANIAMFGAGDEDDTDFLSFS GGARLASLFGLDQAAGHGNEFFGYTAFKQP KKQQTAATGNQATFKTAPATMSTFTILVAT AYHAYYYNGGYVKQKGKGAAVAGHGNEFFGYTAFKQP KKQQTAATGNQATFKTAPATMSTFTILVAT EYRILLYISQQQPVTVARHIVNFELMVEPNNY STFYDDDQRQNWSINFSBEAAAFPFNCQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLPONIVLQGVFDSTANKDKLJLLK LGSGKVIKGWEDMLOMKKGKRLLIVPPA CAVGSEGVIGWTQATTSSLLYFEVEYRVIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPTSIEPFSGGPALTKSNSLSCDLANTSDD AVKAKLISRMAKMGQPBLPILPPQLDSNDSEI EDVNTLQGGOQPVTTSVGPSLQFAHPALPQ MTSQAPQFSVTOLGAFSAALMQVSSLDSISA VSGNAQSFQFYAGMQATAYTQASAVTSQLQ PVRLLYPALSQPHPCSGGMASFLIMTEAR QHINTERMAVSKVADKMDHLIMTKVEELQKH MSGAAPQSVTOLGAFSAALMQVSSLDSISA VSGNAQSFQFYAGMQATAYTQASAVTSQLQ PVRLLYPALSQPHPCSGGMASFLIMTEAR QHINTERMAVSKVADKMDHLIMTKVEELQKH SAGNSMLIPSNSVTINGFSHAALMERSLLGET SQQQSKFLSEKQNRKQLELERNQRYYPGS NLMMEKKNNSLQTATENTOARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQTIXVQAKLSELQET SQQQSKFLSEKQNRKQLELERNQRYYPGS NLMMEKKNSLQTATENTOARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQTIXVQAKLSELQET SGQQSKFLSEKQNRKQLELRNTSQNSCAAAS PPSEKVKLIMNQVPGSLREEFELESYNGRT LGTIMMTITIKMTTQ LQLAQEQEKESSESEEE EKAERRREPSOGGSASSSGOPQARMENEP ESPMYPSOVVELVALQOPGKSESSESEEE EKAERREPSOGGSASSSGOPQARMENEP ESPMYPSOVVELVALQPQESTRLSTSUPPEE GDPLALGPESSGEPOPQLKKDDVTSTIGPHK LGDTLAGGDSGEPGPPQLKKDDVTSTIGPHK LGDTLAGGEPGEPGDFQLGKTDLATTSQDFGHR RGDSEAGLASENGGPGPRINGLSCHGEDEDEVSMKGGR ELEKGATLKAARPKAQQPEEDEDEVSMKGGR PFTILFGDDDDDDDDGGBEDEDEVSMKGGR PFTILFGDDDDDDDDDGUG	1	ł	l	ł		ì	
FIPSIPDIFLSGNKLVTLPKILNLTANLIHLENS LENLDL YFLLKYPHLQILLNOPKISSCSGNQ TYSENPSLEQLFLGENMQLAWETELCWDVF EGLSHLQVLYLNINYLNSLPPOYSSILTALE GLSLNSNRLTVLSHNDLFANLELDISRNQLL APPPDFVSLSVLDTIFNKFICESTSTINWL APPPDFVSLSVLDTIFNKFICESTSTINWL NHTNVTIAGPADDYCVYPDSLSGVSLFSLSTE GCDEEEVLKSLKSKFIPVCVTALLHALTUT TKFRGFCICYKTAQRLVFKDHPQGTEPDMY KYDAYLCFSSKDFTWQNALLHALDTDYSD QNRPILCTEERDPVPGENPANIQDAINSIN KIVCLYSRHTLRGWCLEAFSYAGGRCLSDI NSALIMVVVGSLSQYQLMKHQSIRGFVQKKQ YLRWPEDLQDVGWFHKLSQQILKKEKKKK KDNNPLQTVATIS RDRAGVFPAGKQHAAAAFYDVGGDRPWDS GNTQLPFRNPVKANAMFGAGDEDDTDFLSFS GGARLASLFGLDQAAAGHGNEFFQYTAFKY KKQGTAATGNQATFKTAATATSTTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTI EYRLLIVISQQQPVTVARHVMFEHMYRNNY STFYDDQRQNSIMFESSKAAVEFNQVCICA KCNSTSSLDAVLSQDLIVADGFAVVCGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAYGSEGVIGWTQATDSLLVFEVEVERVKIA KDSGSDGHSVSSBDSAAPSPIPGDLDSDSEL EDVNTLQGGGPVVTTSVQPSLQPAHPALPQ MTSQAAPGSVTQLQAFSALMWGSLDSHSA VSGNAQSRQPYAGMQAYAYQASAVTSQLL PVEPLYPAPLSQPPFIFGGSGDMASFLMFLEAR QHNTEIRMAVSKVADKMDHLMTKVEELQD PVEPLYPAPLSQPPFIFGGSGDMASFLMFLEAR QHNTEIRMAVSKVADKMDHLMTKVEELQBE LKGELEKSNRIEEQNDKISELIEENQRYVSG NLMMEKRNNSLOLQAFSALMWGSVEAAAS QHATEIRMAVSKVADKMDHLMTKVEELQBE SQAQSKFKSEKONRKQLELKVTSLEELITDL RVEKESLEKNLEERNKESAQPRAEIDEI RKSYQEELDKLRQLLKKTRYSTDQAAAGG LLAQA QITALITCONSGMKCEHLAVASAGEHDEI RKSYQEELDKLRQLLKKTRYSTDQAAAGG DPSEKYKKMAQVPQPQLESSKSEELD RKSYQEELDKLRQLLKKTRYSTDQAAAGG LVAGELGYNGKEELSAASSMGVYAAAA QITALITCONSGMKELEKNTSLORGIV EGSPANFSEGVALAGA NGCHEEKSTREFELEESVAGRI LGTMNTIGMATLQLLNQQCKERESSSEELE ERAERRERSGSSAASSGPQAPLINEREP ESPMYPSEGVVERAVPLOQCRERESSSEELE ERAERRERSGSSAASSGPQAPLINEREP ESPMYPSEGVVERAVPLOQCRERESSSEELE ERAERRERSGSSAASSGPQAPLINEREP ESPMYPSEGVVERAVPLOQCRERESSSEELE ERAERRERSGSSAASSGPQAPLINEREP ESPMYPSEGVVERAVPLOQCRERESSSEELE ERAERRERSGSSAASSGPQAPLINEREP ESPMYPSEGVVERAVPLOQCRERESSSESCEL ERSPERPERSGSSAASSGPQAPLINEREP ESPMYPSEGVVERAVPLOQCRERESSSESCEL ERSPERPERSGSSAASSGPQAPLINEREP ESPMYPSEGVVERAVPLOQCPLESSTSTLSTSDPEE GDPLALGPESPGETOPPQLKKDDVTSSTGMH KUDSCS		į.					NLQVLNLSYNLLGELYSSNFYGLPKVAYIDL
LENLDILYFLLRYPHLQEILINQNRFSSSCORD TPSEPPSLEGILFORMIQLA WFTELCWDVF EGLSHLQVLYLNINYINSLPFQVPSHLTAL GLSHNSNLTVALSPROVESHLTAL GLSHNSNLTVALSPROVESHLTAL GLSHNSNLTVALSPROVESHLTAL GLSHNSNLTVALSPROVESHLTAL GLSHNSNLTVALSPROVESHLTAL GLSHSNRLTVALSPROVESHLTAL GLSEFEVLKSLKESLFIVCTVTLTLFLMTHLTV TKRGGCTICYXTAQRLVFXDHDYVDYPSLSSGVLFSLSTE GCDEEFVLKSLKESLFIVCTVTLTLFLMTHLTV TKRGGCTICYXTAQRLVFXDHDYVDYSD QNRPNLCFEERDFVQGNRTANIQDALWAN KYDAYLCFSKDFTWVQNALLKHLDTQYSD QNRPNLCFEERDFVQGNRTANIQDALWAN KYDAYLCFSKDFTWVQNALLKHLDTQYSD QNRPNLCFEERDFVQGNRTANIQDALWAN KYDAYLCFSKDFTWVQNALLKHLDTQYSD QNRPNLCFEERDFVQGNACHLSHLDTQYSD QNRPNLCFEERDFVQGNACHLSHLDTQYSD QNRPNLCFEERDFVQGNACHLSHLDTQYSD QNRPNLCFEERDFVQGNACHLSHLDTQYSD QNRPNLCFEERDFVQGNACHLSHLDTQYSD QNRPNLCFEERDFVQGNACHLSHLDTQYST QNRPNVKANAMFQAODEDDTDTLST GNRQGTAATGNQATPKTAPATMSTFITLVAT AVYBAYNTMGQVYCQKGKGAAVLGNHTTT EYRLLLYISQQQPVTVARHVVPELMVRPNV STFYDDQRQNWSIMFSSKAAVLGNHTAT AVYBAYNTMGQVYCQKGKGAVLGNHTAT EYRLLLYISQQQPVTVARHVVPELMVRPNV STFYDDQRQNWSIMFSSKAAVLGNHTAT EYRLLLYISQQQPVTVARHVVPELMVRPNV STFYDDQRQNWSIMFSSKAAVLGNHTAT AVYBAYNTMGQVTXQGKGGKLLLVVPA CAVGSGGVGGVAVTAGTSLTLVPCVARVLAK CASGGVGGGVAVTAGTSLTLVPCVARVLAK CASGGTGGVAVTSVGQKGKGLLLVVPA CAVGSGGTGGVAVTSVGVGCKGGALLANFPV VSPTSTSPFKSGGPLARTKNSISLGANSVGKLLANFPV VSPTSTSPFKSGGPLARTKNSISLGANSVGKALANFPV VSPTSTSPFKSGGPLARTKNSISLGANSVGKALANFPV VSPTSTSPFKSGGPLARTKNSISLGANSVGKALANFPV VSPTSTSPFKSGGPLARTKNSISLGANSVGKALANFPV VSPTSTSPFKSGGPLARTKNSISLGANSVGKALANFPV VSPTSTSPFKSGGPPFFGGSGGMASSLGHAGGVAAVA VSGNAQSTQPYAGAVTAYQASAVTSOLC PVRLYPAPLSQPPFFGGSGGMASSLGHAGGVAAVA VSGNAQSTQPYAGAVTAYQASAVTSOLC VSQLAGAVAYQACVATAYQASAVTSOLC QNCATSALGANSVGKAAAA QNTATSALGANSVGKAAAA QNTATSALGANSVGKAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAAA QNTATSALGANSVGKAAAAA QNTATSALGANSVGKAAAAAA QNTA	ļ	1	İ				
TPSENPSLEQLFLGENMIQLAWTELCWIDTE GGLSH,QVI,LINIVINSLPGOVPSHITALR GLSLNSNRLTVLSHNDLBANLELDISRNOLL APPPDFVSLSLVDTHRKFLGERESTRIWL HTTNVTHAGPPADDYCVYDDSLSGVSLFSLSTE GCDEEVLKSLKFSLFTVCTVTLTLFHLTTV TKRGFCTCYKTAQRLVFKDHPQGTEPDMY KYDAYLCFSSKDFTWQNALLKHLDTQYSD QNRFNLCFEERDFVFGENRPANIQDAIWNSD KYDAYLCFSSKDFTWQNALLKHLDTQYSD QNRFNLCFEERDFVFGENRPANIQDAIWNSD KIVCLVSRHFLRDGWCLBAFSYAQGKCLSDL NSALIMVVVGSLSQYQLMKHGSIRGFVQKQQ YLRWFEDLQDVGWFLHKLSQQLKKEKEK KDNNIPLQTVATIS 816 2166 A 6646 1 3811 RDRAGWRAGKGHAAAAFYDVGGDRPWD KGQGTAATGNOATFKTAPATHATS GROWNANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAFKY KGQTAATGNOATFKTAPATHAVATNY STYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDLVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDLVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDLVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDLVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDLVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDLVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDVVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQDVVAGHTWATNY STYYDDQRQNWSIMFSSEKAAVEFNKQVCIA KCNSTSSLDAVLSQQVVTAGNICHALTEN LOSGKVKGWBDOMLGKKGGKRLLIVYRS CAVGRGWYTSVYGPSLAMTSDL AVKAKLISRMAKMQQPHLPIPQDLDSSDSLAMTSDL AVKAKLISRMAKMQQPHLPIPQDLDSSDSLAMTSDL AVKAKLISRMAKMQQPHDLPPQLDSSDSLAMTSDL PVRPLYPAPLSQPPFGGSSDMASFLMTEAR QHNTERMAVSKVADKMDH.MT KEELQKH SAGNSALIPSMSVTMETSMMSNIGRIDGEKA QHNTERMAVSKVADKMDH.MT KEELQKH SAGNSALIPSMSVTMETSMMSNIGRIDGEKA QHNTERMAVSKVADKMDH.MT KEELQKH SAGNSALIPSMSVTMETSMMSNIGRIDGEKA QHTALTRUBAGQLTKVAKEELQKYVSDS NLAMCEKNNSLQTATENTQARVLHAEQEKA KVTELAAATAQVSHLQLKKTRVSTDQAAAEQL QUTALITRQNEGPIRKELEKNSKQRAAEGEIDEI RKSYQEELDKLRGLLKKTRVSTDQAAAEQL QUTALITRQNEGPIRKELEKNSKQRAAAS DPSEKVKKIMNQVFQSLKREFELEESYNGVG AQGRAVACKEHLLASKAGELGESSSSEEEE EKAERPRRFSQGQSAASSGQOQAPLINER ESPAYSGQVCRACLALAPPD TSSPPPIGPGSPORGVSCGSAAASSGQPOAPLINER SGDSAAALSEKINGNSSUSGAAAA DPSEKVKIMNQVFQSLKREFELEESYNGGUF ELSTAGSTVAGAALRSHISHGSRSLSGDEE EKAERPRRFSGGCSAAARSGCPOAPLINER ELSTAGSTVAGAALRSHISHGSRSLSGDEE DELFKGATLKALRPRAQPEEDED	1	1					
EGLSHLQVLÝININYINSI.PFGVPSHLTALG GLSINNRILTVIS.PFOVPSHLTALG GLSINNRILTVIS.PFOLPSNILDISRNQLI APPPDVPSLSVLDITINKFICECEL.STSININ HTINVTIAGPPADYCVYPDSLSGVSLFSLSTE GCDEEVILSLKFSLFIVCTVITLIT.HMTILTV TKPRGFCICYKTAGRLVFKDHPQGTEPDMY KYDAYLCFSKGDFTWQNALLHDTYSSD QNRFNLCFERDPVPGENRPANIQDAINYIS GNRFNLCFERDPVPGENRPANIQDAINYIS KYVCLVSRHHLRDGWCLEAFSYAGGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGFVQWG YLRWPEDLQDVGWFLHKIS.QQILKKEKEKK KDNNIPLOTIVATIS BRAGVEPAGKQHAAAAPYDVGGDRPWDS GNTQLPPRNPVKANAMFGADBEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KKQQTAATGNQATPKTAPATNSTPTILVAT AVHAYRYTINGQYVKQGKFGAAVLGNHTTI EYRLLITSQQQPVTVARHNYBELMVRPNNY STFYDDQRQNWSIMFESSKAAVEFNQVCICA KCNSTSSLDAVLSQDLIVADGPAVVGOSLE VAYTGWLFQNHVLGQVFDSTANKDKLLIKL LGSGKVIKGWEBGMLGMKKGGKRLLIVPPA CAYGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSBDAAPSPIPQDLSSDSEL EDWNTLQGGGQPVVTFSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSISSE AVKAKLISRMAKMGGMPLEIPPQLDSNDSSEL EDWNTLQGGGQPVVTFSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSISSE AVKAKLISRMAKMGGMPLEIPPQLDSNDSSEL EDWNTLGGGGPVVTFSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSISSE AVGNOAGSPGPYAGMQATATYQASAVTSQLQ PVRELYPAPLSQPFHFQGSSDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMKTKVEELQKH SAGNSMLPSMSVTMETSIMMSNIQRIQDENE LKGELEKSNRIEGPONKISELERNQRYVEGS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQKKSTEL LQQLTESLKETDLLRGQLTKVQARLSELDTI LGCMANSAVSAVGLANGAVGVAAAS DFSEKVKKIMNQVFQSLLKRITKVSTDQAAAEQIG PVELSSELEKNLSERKKSAQERSQAEEDIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQIQ QTALITKQNEQHEKLEKNKSQNGSVEAAAS DFSEKVKKIMNQVFQSLRRFFELEESYNGVE EGAGERPRRFSQGOSASASSGOPQAPLNRERP ESPMYPSGQVVCRSAVAPDPPLQESTRLSLTSUPPE GDPLALGFESPGEPQPPQLKKDDHTRSTD PSEKVEKINNQVFQSLRRFFELEESYNGVE EEKAEERPRRFSGGOSASASSGOPQAPLNRERP ESPMYPSGQVVCRSVAPDPPQLSCTSHAVLOPD TSPPPPTLGFGDDDDDDBDUGC BCLAGFTRAFAQFEEDEDEDEVSNKKGR PPTPLFGGDDDDDDDDDDUDUGC BCLAGFTRAFAQFEEDEDEDEVSNKKGR PPTPLFGGDDDDDDDDDDUDUG		J	}]	
GLSLNSNRLTVLSHNDLANLELDISRNOLL APPPDYFVSLSLDTIFNKFICECELSTFINWL NHTINVTIAGPADDYCVYPDSLSGVSLFSLSTE GCDEERVLKSLKSFLFIVCTVTLHUTTUT TKFRGFCTICYKTAGRLVFKDHPQGTEPDMY KYDAYLCFSSKDFTWQNALLKHLDTQYSD QNRFNLCEERDFVFGENRPANIQDAIWNSR KIVCLVSRHFLRDGWCDAIRGFVQNQL NSALIMVVVGSLSQYQLMKHOSPIGFVQNQQ YLRWFEDLQDVGWTLHKLSQQLKKEKEK KDNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGYRPAGKQHAAAAFYDVGGDRPWG YLRWFEDLQDVGWTLHKLSQQLKKEKEK KDNIPLQTVATIS GNTQLFPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAATTGAAATCHAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAATCHAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAATCHAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAATCHAAAFYDVGGDRPWG KGGGTAATGNAAFTGAAATCHAAAFYDVGGDRPWG KGGGTAATGNAATTGAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAATCHAAAATCHAAATCHAAAATCHAAAAAAATCHAAAAAAAA	1						1
APAPDVYSLSVLDITHIKRICECELSTRINWL NHITNYTAGPADITYCVYPOSISGYSLFISLTS GCDEERVLSILKFSI FVCTVTLTILTIMIT TV TKFRGFCFCTKTAGRL VRKDHPQGTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTQYSD QNRFNLCFERDFVPGENRFANIQDALMYSR KIVCLVSRIFLERDGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSQYQLMKHGSRIGFVQKQQ YLRWPEDLQDVGWPLHKLSQQLKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAAMFQAGDEDDTDLSPS GGALASLEGLDQAAAGHOFQYTAPKQP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGHTTR EYRLL1YSQQQVTVARIHVFELMWRPNNY STTYDDQRQNWSIMFSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLIRLK LGSGCVIGWTQATDSILVFEVEVRRVIA KDSGSDGHSVSSRDSAAPSIPGADNLSADPV VSPPTSIPFKSGGPALRTKSNSLSEQLAINTSPD AVKAKLISRMAMGQPMPLPQLDSNDSSE EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPTAGMQAYATYQASAVTSQLQ PVRPLYPAPLSQPHHPQGSGDMASTLMHEAR QHNTERMAVSKVADKMDHLMKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIQCHER LKQELLEKSNEEQONDISELENGRYVEQS NLMMEKRNNSLQTATENTQARVLAEGEKA KVTELAAATAQVSHLQLKMTALEECRY LKQELLEKSNEEQONDISELENGRYYEQS NLMMEKRNNSLQTATENTQARVLAEGEKA KVTELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLIRGQLTKVQAKLSELQET SEQAQSKYKSKRAQVSHCLAAGQLS LVQAELQTYQWAAKCEHLLASAKDEHLQQY CVCAQRDAYQKLVQLQEKSVCFAALAQA QITALTKONEQOKKSCHAAAS DPSEKVKRMNQVFQSLREFFELEESYNGRTI LGTIMMTIKMVTQLLNQQECKEESSSEEEE EKAEERPRRPSQCQSASASSGGPQAPLNERP ESSPMYSEQVVERAVPLPQQLTTSQDGHRR KGDSAAALSEINDGSLPPELSCPSRFVLGPP TSIPPPELGPVDRAGSCECESLASHAARQPUN PSGKVCVVERVAPPGQLQESSTRLSLTSUPPE GGDPAALGPESPGEPQPPQLKXDDVTSSTPHE ELSSTEAGSTVAGAARLSPKAQPEEDEDEVSMKGR PPPTPLFGDDDDDDDWLG ELEFKGATLKALRYKAQPEEDEDEVSMKKGR PPPTPLFGDDDDDDDWLG						!	1 · · · · · · · · · · · · · · · ·
NHTHVTTAGPPADIYCVYPDSLSGVSLSISLSTSISTE GCDEEPULSLKYSI FIVCTUTLIFLMTILTY TKPRGFCFICYKTAQRLVFKDHPQGTEPDMY KYDAYLCPSSKDFTWVQNALLKHLDTQYSD QNRPNLCFSERDFYPGDRNFANIQDAUMYSR KIYCLVSRIFLRDGWCLEAFSYAQGRCLSDL NSALMVVVOSSLSQYQLMKGRGFVQKQQ YLRWFEDLQDVGWFLHKLSQQLLKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAFTDVGGDRPWDS GNTQLPPRNPYKANAMPGDEDDTDFLSPS GGARLASLFGLDQAAAGHONEFFQYTAPKQP KKQQTAATONQATAFKTAPATMSTPTILVAT AVHAYRXTNGQYVKQGFGAAVLGNHTIT EYRLLYISQQQPVTVARHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAPFNKQVCLA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGOVPDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKBLLIVPPA CAVGSEGVIGWTQATDSLTVEVRRVKLA KDGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQYMLPIPPQLDSNDSEL EDWNTLQGGQPVVTTSVQPSLQPAHPALPQ MTSQAPGPSYTQLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHPGSGDMASFLIMFAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSNSVTMETSMRNGRUIGDERE LKQELEKSNRIEEQNDKISELIERNQRYVEQS NLAMEKRNNSLQTATENTQARVLAEQKA KVTEELAAATAQVSHQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKYKESKQNRKQLELKYSLEELDTL RVEKESLEKNLSERKKRSAQERSQAFEDIDEI RKSYQEELDKLRQLLKTKYSTOQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQOYQ EVCAQRDAYQQKLVQLQEKSESSEEEEE EKAEERPRRPSQEQSASASSQQPQAPLNBERP ESSMYSEGVYGALKRQLKKTRSTOQAAAEQLS DPSEKYKLMMQVFQSLRGEFELEESYNGRTI LGTIMNTIKMYTQLLLNQGCKEEESSSEEEEE EKAEERPRRPSQEQSASASSQQPQAPLNBERP ESSMYSEGVYGALKRQLKTRYSTOQAAAEQLS DPSEKYKLMMQVFQSLRGEFELEESYNGRTI LGTIMNTIKMYTQLLLNQGCKEEESSSEEEEE EKAEERPRRPSGEQSASASSQQPQAPLNBERP ESSMYSEGVYGAVLATISQDGHRR KGDSEAALSEIKDGSLPPELSCIPSIRKYLGPP TSIPPPLGYWDSGCEEGLAASPMAAKPDM PSGKVCVREVAPDGPQCKSCESSSEEEEE EKAEERPRRPSGEQSASASSQQPQAPLANDREP ESSFRYPSEGPOPPQLATTSQDGHRR KGDSEAALSEIKDGSLPPELSCIPSIRKYLGPP TSIPPPLGYWDSGCEEGLAASPMAAKPDM PSGKVCVREVAPDGPQLQSSTRLSLTSUPPEE GDPLALGPESSFEREPEDEDEVSMKKGR PPPTPLRGDDDDDDDWLG ELSSTAAGSTLAAARAKARGRE	1		ĺ			1	I
GCDEEVLKSILKPSI.FVCTVTI.II.TIMIT.V TKERGFCTCVKTAQRI.VFKDIPQGTEPDMY KYDAYLCFSSKDFTWVQNALI.KHLDTQYSD QNRFNLCFEERDFVGENRFANIQDAINVSR KIVCLVSRIFILERDGWCLEAFSYAQGRCI.SDI. NSALIMVVVGSI.SQYQLMKHQSIIGGFVQKQQ YLRWFEDLQDVGWFLHKI.SQQLKKEKEKK KDNNIPLQTVATIS BDRAGVERAGKQHAAAFYDVGGDRFWDS GGABLASLFGI.DQAAAGHGNEFQYTAFKQP KKGQGTAATIGNQATFKTAFATMSTFTIL.VAT AVHAYRYTNGQYVKQGKFGAAVLGNHTIR EYRIL.YISQQQYTVARHIVFELMVRPINY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLI.VAQFAVEVGDSIE VAYTGWLPQNHVL.GQVFDSTANKDKLLRLK LGSGKVIKGWEDDMIMKKGGKRLI.PPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNI.SADPV VSPPTSIPKSGEPALRIKNSNI.SEQLAINTSPD AVKAKLISRMAKMGQFMLPILPPQL.DSNDSEI EDVNTLQGGGQPVVTPSVPSLQPAIPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTMIEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSNSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NIAMMEKRNISQTATETNJAKLSQLELKVTSLEEELIDL RVERESLEKNLSRKKSAQERSQAAEDEL RKSYQEELDKLNGLIKKTKYSTDQAAAEGLD RKSYQEELDKLNGLIKKTKYSTDQAAAEGLD RKSYQEELDKLNGLIKLKTKSTDQAAEGLD RKSYQEELDKLNGLIKLKTKSTDQAAEGLD RKSYQEELDKLNGLIKLKTKSTDQAAEGLD RKSYQEELDKLNGLIKLKTKSTDQAAEGLD RKSYQEELDKLNGLIKLKTKSTDQAAEGLD RKSYQEELGROWLOVEGAVELOVICA QTTALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKLMONVOPOSLRREPLESYNGRTI LGTIMNTIKMVTLQLLNQEQEKEESSSEEEE EKAEERPRPSGEQSASASSOQPQAPLNERRP ESSPMYPSGVVCEAVLPQPQALTTSQDGHER KGDSEAALSEIKDGSLPPELSCIPSIRKTLOPP TISPPPLGPYBMSCECTESLAGSTNAAKPDD PSGKVCKUNQVCQSLREPLSSYNGRTI LGTIMNTIKMVTLQLLNQEQEKEESSSEEEE EKAEERPRPSGEORSASSSOQPQAPLNERRP ESSPMYPSGVVCEAVLPQPALTTSQDGHER KGDSEAALSEIKDGSLPPELSCIPSRINGRTI LGTIMNTIKMVTLQLLNQEGEKEESSSEEEE EKAEERPRPSGEORSTELSSYNGRTI LGTIMNTIKMVTLQLLNQEGEKEESSSEEEE EKAEERPRPSGEORSASASSOQPQAPLNERRP ESSPMYPSGVVCEAVLPQPALTTSQDGHER KGDSEAALSEIKDGSLPPELSCIPSRINGRTI LGTIMNTIKMVTLQLLNQEGEKEESSSEEEE EKAEERPRPSGEORSTELSSYNGRTI LGTIMNTIKMVTLQLLNQEGEKEESSSEEEE EKAEERPRPSGEORSTELSSTORGSTELSSTORGS PPPTPLFGDDDDDDDWLG		ł					
TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY KYDAYLCFSERDFTWYQMLIAHLDTDYSD QMRFNLCFEERDFTWGENRPANIQDAIWNSR KIVCLVSHIFLRDGWCLEAFSTAQGRCLSDL NSALIMVVVGSLSQYQLMKHGSIRGFVQKQQ YLRWPEDLQDVGWPHIKLSQQLKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGGDRFWDS GMALAALSFGLDQAAAGHGNEFFQYTAPKQP KKGQOTAATGMQATPKTAPKQP KKGQOTAATGMQATPKTAPKQP KKGQOTAATGMQATPKTAPKTPILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTIT EYRIL1YISQQQPVTVARIHVNFELMWYRNNY STIYDDQRQNWSIMFESSKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVPDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIPPPA CAVGSEGVIGWTQATDSLIVPPA CAVGSEGVIGWTQATDSLIVPPA CAVGSEGVIGWTQATDSLIVPPQLDSHDSEI EDVNTLQGGQPVYTTSVQPSLQPAHPALPQ MTSQAPQPSYTOLQAPSAAMSFLIMFEAR VSGNAQSFQPYAGMQAYAYTQASAVTSQLQ PVRPLYPAPLSQPPHPQGSGNASILMFEAR QHNTEIRMAVSKVADKMDHLMTKVELQKH SAGNSMLPSNSVTMETSMIMSNIQRIQDENER LKQELEKSNRIEGONDKISELERNQRYVEQS NLAMMEKRINSLQTATENTQARVLHAEQEKA KVTELAAATAQVSHQLKWTSLQEELDEI LKQELEKSNRIEGONDKISELERNQRYVEQS NLAMMEKRINSLQTATENTQARVLHAEQEKA KVTELAAATAQVSHQLKWTSLGELDEI RKSYQEELDKLKVSLLEELDID RVEKESLEKNLSERKKSAQERSQAAEGLOT RKSYQEELDKLKVILLEELDID RVEKESLEKNLSERKKSAQERSQAEDIDEI RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRVILLYKTSTDQAAEGLOE RKSYQEELDKLRAGRAAKPDN PSGKVCVGVEAVPLPQGLQEKEESSSEEEEE EKAEERPRPSGQSASASSGQPQAPLNREPP ESSPMYSSQVVEAVPLPQGLQEKEESSSEEEEE EKAEERPRPSGQSASASSGQPQAPLNREPP ESSPMYSSQVVEAVPLPQGLQESEESSSEEEEE EKAEERPRPSGQSASASSGQPQAPLANDREPP ESSPMYSSQVVEAVPLPQGLASSTRLS.LTSUPBE GDPLALGPESSPEGPPOPPQLXSTDGTHK ELLSTEAGSTVAGAALRSHHSQRSSLSGDEE DELFKGATLKALRRKAAQPEEEDDEVSMKG	1	}	!	1	•	ĺ	L
KYDAYLCESSKDFTWQNALLKHLDTQYSD			ŀ				
ONEPHLCPERDPYPGENRPANIODAIWNSIR KIVCL VSRIPLRDGWCLESYVAQGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ YLRWPEDLQDVGWFLHKLSQQILKKEKEKK KDNNPIQTVATIS BIFAGYRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRIPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KKGQCTAATGNQATFKTAPATMSPTTLVAT AVHAYRYTNGQYVKQCKFGAAVLGNHTITR EYRLLYISQQQPVTVARIHVNPELHVRPNNY STYPDDQRONWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLPQNHVSIGPSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLPQNHVIGQFSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLPQNHVIGQFSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLPQNHVIGQFSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLPQNHVIGQFAVEVGDSLE VAYTGWLPQNHVIGQFAVEVGDSLE VAYTGWLPQNHVIGQFAVEVGDSLE VAYTGWLPQNHVIGQFAVEVGDSLE VAYTGWLPQNHVIGQFAVEVGDSLE VAYTGWLPQNHVIGQFAVEVGDSLE LGSGKVIKGWEDGMLGMKKGKRLLIVPPA CAVGSEGVIGWTQATDSLIVFEVERVKIA KDSSGDGHSVSSRDSAASPSPPGADNLSADPV VSPPTSIFFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPPQDLDSNDSEL EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MITSQAPQFSVTGLQAPSAALMGVSSLDSHSA VSGNAQSFQPYAGMQAVAYPQASAVTSQLQ PVRRLYPALSQPHFGQSGDMASFILMEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIQENER LKQELLKSNRIEEQNDKILELERNGSYVEGS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKNTAIQKKETEL QMQLTESLKETDLLRGQLIKKVGAKLSELQET SEQAQSKFKSEKQNRKQLELKVTISEELITDL QWQLTESLKENLSERNKKSAQERSQAEEEIDEI RKSYQEELDKLRGLIKKSAGCRSQAEEEIDEI RKSYQEELDKLRGLIKKSAGCRSQAEEEIDEI RKSYQEELDKLRGLIKKSAGCRSQAEEEIDEI RKSYQEELDKLRGLIKKSAGCRSQAEEEIDEI RKSYQEELDKLRGLIKKSAGCRSQAEEEIDEI RKSYQEELDKLRGLIKKSAGCRSQAEEEIDEI RKSYQEELDKLRGLIKKSAGCRSQAEERIDEI RKSYQEELDKLRGLIKKTRYSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDGNSVEAAAS DPSEKVKKIMOVPGSLKKSAGCRSQAEERIDEI RKSYQEELDKLRGLIKKSGGRSVEAAG DPSEKVKKIMOVPGSLKREFFELESSYNGET LGTIMNTIKMVTLQLLNQCQCKEESSSIEEEE EKAEERPRRPSQGSASASGQPQAPLNRERP ESPMYPSQVVEEAVPLPPQALTISQDGHRR RGDBAALASEIKDOSLPFELSCIPSRIVLGPP TSTPPPELGPSMDSCCESSLASSRYMAAKPDN PSGKAVCVREVAPDOPLQESSTELSLTSDPEE GDPLALGPESPGEOPPOPLQESSTRLSLTSDPEE GDPLALGPESPGEOPPOPLQESSTRLSLTSDPEE GDPLALGPESPGEOPPOPLQESSTRLSLTSDPEE GDPLALGPE			1				l
KIVCLVSRHFLRDGWCLEAFSYAQGCLSDL NSALMWVVGSLSQVQLMKHOSIRGFVQKQQ YLRWFEDLQDVGWFLHKLSQQLKKEKEKK KDNNIPLQTVATIS RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMGAGDEDDTDLSPS GGARLASI-GLDQAAAGHGNEFFQYTAPKQP KKGQGTAATGNQATPKTAPATMSTPTLVAT AWHAYRYTNGQVVKQGKGAAVLGNHTTR EYRILLYISQQPVTVARHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCLA KCNSTSSLDAVLSQDLIVJADGPAVEVGGSLE VAYTGWLPQNHVLGQVFDSTANKDKLIRLK LGSGKVIKG WEDGMLGMKGGRALLIVPPA CAVGSEGYIGWTQATDSILVFEVEVRRVKLA KDSGSGGHSWSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQ-DSNDSEI EDVNTLQGGGPVVTFSQSGDMASFLMTEAR QHNTEIRMAVSKVAZKMPYQASAVTSQLQ PVRRLYPAPLSQPPIFGQSGDMASFLMTEAR QHNTEIRMAVSKVAZKMPAYQASAVTSQLQ PVRRLYPAPLSQPPIFGQSGDMASFLMTEAR QHNTEIRMAVSKVAZKMRNSQTYAGSAV SGRAQSFQYAGMQAYAYLAGKSELERNGYVEQS NLMMEKRINSLQTATENTQARVLHAGQEKA KVTEELAAATAQVSHLQLKNTTAHCKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSTQEELDRLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLGSVCFAAAS DPSEKVKKIMOVPGSLESVCFACLALQA QITALTKQNEQHIKBLEKNRSQMSGVEAAAS DPSEKVKKIMOVPGSLESVCFACLALQA QITALTKQNEQHIKBLEKNRSQMSGVEAAAS DPSEKVKKIMOVPGSLESVCFACLALQA PSPSELONLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYG EVCAQRDAYQQKLVQLQSKYCFACLALQA PSPSELVKKIMOVPGSLESSTELESSEEEE EKAEERPRPSQEQSASASSQQPAAAN DPSEKVKKIMOVPGSLESTCLESSTORET LGTIMMTIKMVTLQLLNQEQEKEESSSEEEE EKAEERPRPSQEQSASASSQQPALNICEP ESPMYPSEQVVELAVYLPQALTISQDGHRR KGDSSAEALSEIKDGSLPFELSCTPSHRVLGPP TSIPPFELGPSMDSCCEESLAASPMAKJPDN PSGKWCVREVAPDOPLQESSTELSTSDPEE GDPLALGPESPGEOPPPQHALTISQDGHRR KGDSSAEALSEIKDGSLPFELSCTPSHRVLGPP TSIPPFELGPSMDSCCEESLAASPMAKJPDN PSGKWCVREVAPDOPLQESSTELSTSDPEE GDPLALGPESPGEOPPPQHALTISQDGHRR KGDSSAEALSEIKDGSLPFELSCTPSHRVLGGPP TSIPPFELGPSMDSCCEESLAASPMAKJPDN PSGKWCVREVAPDOPLQESSTELSTSDPEE GDPLALGPESPGEOPPPQHALTISQDGHRR KGDSSAEALSEIKDOBLDDDDDDDUBUG			1]			
NSALIMVVVGSLSQYQLMKHQSIRGFYQKQQ YLRWPEDLQDVGWPLHKLSQQLKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMFQAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KGQGTAATGNQATTAPATMSTPTLLVAT AVHAYRYTNGQYVKQGKFGAAVLGHITTR EYRILLYISQQPVTVARHVNFELMVRPNNY STFYDDQRQNWSIMFSISKAAVENKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFONHVLGDVFSTANKDKLILKK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVESGEVIGWTQATELYFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPTSIFFKSGEPALRTKSNSLSQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGPVVTPSVQPSLQPAHPALPQ MTSQAPPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQYAGMQAYAYPQASAVTSQLQ PVRLYPAPLSQPHPGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMMSNIQRIQENER LKQELLKSNRIEGONISELIERNQRYYEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKSELEKNLSERKASQERSQAEELDEI RKSYQEELDKLRQLLKKTRAHQKKETEL QMCLTESLKENDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL LGTIMNTIKMVTLQLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFALCALQA QITALTKQNEQHIKELEKNKSQNGSVEAAAS DPSSKVKKIMMQVFQSLREFELEESYNGRTI LGTIMNTIKMVTLQLLAQQQCGKEESSSEEEE EKAEERFRRPSQEGSASSGGPQAPAINERP ESPMVPSEQVVEEAVPLPPQALTTSQDFRR KGDSEAALSEIKDGSLPPELSCFSHRVLGPP TSIPPEPLEGDDDDDDDIGLG	ľ		1				QNRFNLCFEERDFVPGENRP\ANIQDAIWNSR
SIGNA Signa Sign			}				
KDNNIPLQITVATIS			1				1
816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMFQAGDEDDTDFLSPS GOARLASLFGLDQAAAGHGMEFFQYTAPKQP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVGKGAAVLGHHTR EYRILLYISQQQPUTVARIHVMFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMGLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPFTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVYTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSTQPYAGMQAYAYPASASVTSQLQ PVRPLYPAPLSQPHPHQGSGDMASFIMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTAESMMSNIQRIGENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHCLIKVTSLEELETDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLIKQLLKKTRVSTDQAAASQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQGKLVQLQEKSVCPACLALQA QITALTKONPGHIKEI EKNKSQMSGVEAAAS DPSEKVKKIMNQVPQSLRREFELEESYNGRTI LGTIMNTIKMYTLQLLNQQEQCKEGSSSEEE EKAEERFRRPSQEQSASASSGQPQAFLNRERP ESPMYPSEQVVEEAVPLPPQALTISQDGHRR KGDSEAEALSEIKDGSASASSGQPQAFLNRERP ESPMYPSEQVVEEAVPLPPQALTISQDGHRR KGDSEAEALSEIKDGSASASSGQPQAFLNRERP ESPMYPSEQVVEEAVPLPPQALTISQDGHRR KGDSEAEALSEIKDGLESSTEEEL EKAEERFRRPSQEQSASASSGQPQAFLNRERP ESPMYPSEQVVEEAVPLPPQALTISQDGHRR KGDSEAEALSEIKDGLESSTEELSTSUPPE GDFLALGFESSPEEPQPQLKKDDVTSSTOPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPFTPLFGDDDDDDDDDUCHG		İ					
GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAFKQP KGGGTAATGNQATFKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRLLYISQQQDYVARHVMFELMVRPNNY STPYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGFAVEVGDSLE VAYTGWLPQNHVYARHVMFELMKYKDLYLK LGSGKVTKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSBDSAAPSPIPGADNLSADPV VSFPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEL EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNREEQNDKISELERNQRYVEQS NLMMKEKNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKWTSLEGELIEL QMQLTESLKETDLLRQQLTKVQAKLSELQET SEQAQSKFKSERQNRKQLELKVTSLEEELITDL RVEKESLERNLSERKKKSQASQAEEEIDEI RKSYQEELDKLRQKTVAQAKLSELQET SEQAQSKFKSERONRKQLELKVTSLEEELITDL RVEKESLERNLSERKKKSQMSQVEAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQGSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCTPSRTVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKVCVREVAPPOGPLQESSTLSLTSUDPEE GOFLALGFESSEETELSSLTSUDPEE GOFLALGEPSSEFELESSTRSLTSUDPEE GOFLALGFESSEETELSTSTLSSUDPEE GOFLALGFESSEETELSTSTLSSUDPEE GOFLALGFESSEETELSTSTLSSUDPEE GOFLALGFESSEETELSTSTLSSUDPEE GOFLALGFESSEETELDEDDDWLG	<u></u>	-					
GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVYKQGKFGAAVLGNHTITR EYRILLYISQQQPVTVARIHVIFELMVRPINNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVPDSTANKDKLLRLK LGSGKVKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRVKIA KDGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLANTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPIGGSDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIGRIIQENER LQGELLEKSNRIEDENDKISELEERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHQLKMTAHQKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQULKKTRYSTDQAAAEQLS LVQAELQTOWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLUQKESVCFAACALQA QITALTKQNEQHIKELEENKSQMSGVEAAAS DPSEKVKKIMQVFQSLRREFELEESYNGRTI LGTIMNTIKMYLQLLNQQPGKEEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSRIVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYDDN PSGKVCVREVAPDGPLQESSTRLSLTSDDEEE GDPLALGPESSGEETEL ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG	816	2166	A	6646	1	3811	
KKGQGTAATGNQATPKTAPATMSTFTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVIVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSPFKSEGPALRTKSNSLSSQLANTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVOPSLQPAHPALPQ MTSQAPQPSVTQAPSALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPHPQGSGDMASFLMTEAR QHNTERMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NILMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTRVSTLDAKLELETDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYQEELDKLILRGQLTKVAKLSLLQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYGEELDKLILRGQLTKVAKLSTLQVQ EVCAQRDAYQQKLVVQLQEKSVCFAVCLALQA QITALTKQNEQULIKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVVQLQEKSVCFAVCLALQA QITALTKQNEQHIKELEKNSQMSGVEAAAS DPSEKVKKIMNQVFGSLEREFELEESYNGRTI LGTIMNTIKMVTQLLNQQEQEKEFSSSEEIE EKAEERPRRPSQEQSASASSQQPQAPINRERP ESPMVPSEQVVEAVPLPQALTTSQDGHRR KGDSEABALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKVVCVREVAPDGPLQESSTRLSITSUPEE GDPLALGPESSPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHISQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD		İ					GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
AVHAYRYTNÖQYVKQGKFGAAVLGNHTTR EYRILYISQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLIKL LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHISVSGAPSPHPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISMAKMGGPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVOPSLQPAIPALPQ MTSQAPQPSVTQLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPHPGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEELITDL RVEKESLEKNLSERKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQUTQLGKSVCFAKCLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMQVPQSLREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMVYSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESIAASPMAAKYDDN PSGKVCVREVAPDGPLQESSTRLSLTSUDPEE GGPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHISQRSSLSGDEE DELFKGATIKAARPKAQPEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDUNLG							GGARLASLFGLDQAAAGHGNEFFQYTAPKQP
EYRILYISQQPYTVARHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKCKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAMPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTFSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSSQPYAGMQATAYPQASAVTSQLQ PVRRLYPAPLSQPPHPGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKNTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELDEI RKSYQEELDKLRQLKKTSLEELDEI RKSYQEELDKLRQLKKTSLEELDEI RKSYQEELDKLRQLKKTSLEELDEI RKSYQEELDKLRQLKKTSLEALQAY QTAALTKQNEQHIKELEKNSQMSGVEAAAS DPSEKVKKIMQVPQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEAVPLPPQALTTSQDGHRR KGDSEAEALSEKDGSLPPELSCIPSHRVLGPP TSIPPPELGPVSMDSECEESLAASPMAAKPDN PSGKVCVREVAPDGPLQESSTRLS.ITSDPEE GDPLALGPESPGPPQPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHSQRSSLSGDEE DELFKGATIKARPKAQPEEDEDEVSMKGR	Ī						
STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPFTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTTSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMIEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMINSNIQRIGENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRQQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQCRSQAEGEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQVI EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDFHR KGDSEAEALSEIKDGSLPFELSCIPSHRVLGPP TSIPPEPLGPVSMDSECCESLAASPMAAKPDN PSGKVCVREVAPDGPLQESSTRLSLTSUDPEE GDPLALGPESSGEETQFPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDIDWLG	1	l					
KCNSTSSLDAVLSODLIVADOPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTTSVQPSLQPAIFPALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQELLEKSNRIEEQNDKISELIERNQRYVEQS NIMMEKRNSLQTATENTQAVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLITESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRGLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELENKSQMSGVEAAAS DPSEKVKKIMQVPGSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEGVVEEAVPLPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECESLAASPMAAKPDN PSGKNCVCREVAPDGPPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAOPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDWLG	1	1	·				I
VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPFTSIPFKSGEFALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNREEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEELITDL RVEKESLEKNLSSEKKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLREEFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKUDN PSGKVCVREVAPDGPLQESSTRLSLTSUPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRTSHHSQRSSLSGDEE DELFKGATLKALRFKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDVLG			!	'			l
LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVPRLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEQONDKISELERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQOEQEKEESSSEEEE EKAEERPRPSQCQSASSSGQPQAPLNERRP ESPMYPSEQVEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYDN PSGKVCVREVAPDGPLQESSTRLSLTSUPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHASQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	İ			1			
CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQEVVTFSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRINSLQTATENTQARVLHAEQEKA KVTEELAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRSQEGSASASSGOPQAPLNRERP ESPMYPSEQVVEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKIPDN PSGKVCVREVAPDGPLQESSTRLSLTSUPFEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRTSHHSQRSSLSGDEE DELFKGATLKALRFKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	J						
KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTTSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGOPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPFELSCIPSHRVLGFP TSIPPEFLGVSMDSECEESLAASPMAAKVPDN PSGKVCVREVAPDGPLQESSTRLSLTSUPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKAALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD							
VSPPTSIPFKSGEPALRIKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQELLEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEKDGSLFFELSCIPSHRVLGFP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PGSKVVCVRVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPSHASQRSSLSGDEE					,		
AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGMSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHASQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDWLG	1						
EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKIPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR		[
MTSQAPQPSVTĞLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESSPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGFVSMDSECEESLAASPMAAKJPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESSPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR	i	ĺ					
VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHILLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGKVCVREVAPDGPLQESSTRLSLTSUPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDUWLG							
PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKLIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDVLG						•	
QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSYCFA(CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVPQSLRREFELEESYNGRTI LGTIMNTIKMVTIQLLNQEQEKEESSSEEEE EKAERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGKVVCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDVLG	1						
SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAVCLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEKDGSLPPELSCIPSHRVLGPP TSIPPPLGPVSMDSECEESLAASPMAAKYPDN PSGK\VCVREVAPDGPLQESSTTLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD		ļ :)	
LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVAAAS DPSEKVKKIMNQVFQSLREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGK\VCVREVAPDGPLQESSTRLSLTSJDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	1						
NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGKVCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	1						SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER
KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGKIVCVREVAPDGPLQESSTRLSLTSIDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDUWLG						1	LKQEILEKSNRIEEQNDKISELIERNQRYVEQS
QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESFGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDWLG							
SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDUWLG					l	ľ	
RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDUWLG							
RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKVCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDUWLG				[DVENEOU ENTE OFFICE AND A STREET OFFICE AND A
LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\UPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDWLG	1						PVENOPEL DVI POLI VVIDVODDO A A SOCI
EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\UPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDWLG	1		ł	•	l		
QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\UPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDUWLG				1	•		LVQAELQIQWEAKCEHLLASAKDEHLQQYQ
DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDWLG							EVLAQKDA YQQKL VQLQEKSVCFA\CLALQA
LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG				ĺ			
EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDUWLG			}	J			DESCRIPTION OF STREET
ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDUWLG				ŀ			LUILMNIKMVILQLLNQQEQEKEESSSEEEE
KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDUWLG							LKALLKPKKPSQEQSASASSGQPQAPLNRERP
TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDUWLG			İ				
PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDUWLG				ŀ			
GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDUWLG			ł	į	ł		
ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG							
DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG		.[}	İ	.]	•	
PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	.		1				
	, ,]	j	J	j	J	
61/ 216/ A 6649 63 1073 FFRSSSDNGSPIROYE/HSTPAHOGPVMGLEG	F	2162		77.5			
	61/	2167	A	6649	63	1073	FFKSSSDNGSPIRQYE/HSTPAHQGPVMGLEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK
						VFHSGTAAKSITKKCEKRSSSWKETELVVVD TPGIFDTEVPNAETSKEIRCILLTSPGPHALLL VVPLGRYTEEEHKATEKILKMFGERARSFMIL IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG DRYCALNNKATGAEQEAQRAQLLGLIQRVV RENKEGCYTNRMYQRAEEEIQKQTQAMQEL HRVELEREKARIREEYEEKIRKLEDKVEQEKR KKQMEKKLAEQEAHYAVRQQRARTEVESKD GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP GISARGLSHEGRKQLAVNLTRVLALYRSILDA YIIEFFYTDNLWDTLPCSWQSALDGLKPPQLA TMLLGMPGEGEVVRYRSVWPLTLLALKSTA CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPGNHLSRFMALGLGLMVKSIEGDQRL VERAQRLDQELLQALEKEEKRNPQVVQTSPR HSPHHVVRWVDPTALCEELLLPLENPCQGRA RLLLTGLHACG\DLSVALLRHFSCCPEVVALA SVGCCYMKLSDPGGYPLSQWVAGLPGYELP YRLREGACHALEEYAERLQKAGPGLRTHCY RAALETVIRRARPELRRPGVQGIPRVHELKIEE YVQRGLQRVGLDPQLPLNLAALQAHLAQEN RVVAFFSLALLLAPLVETLILLDRLLYLQEQA LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG QALSVLETEDS
819	2169	A	6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLLA AAGTAVGDRCERNEFQCQDGKCISYKWVCD GSAECQDGSDESQETCLSVTCKSGDFSCGGR VNRCIPQFWRCDGQVDCDDGSDEQGCPPKTC SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD CCEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLARNDMRSCLTEGYEAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAGRGNEKKPSSVRALSIVL PIVLL VFLCLGVFLL WKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI EGYLHILLISKAEDYRSYRFPKLTVITEYLLLF RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	-	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isolaucine, K=Lysine, L=Laucine,
seq-	uence]	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
LULINCE .	1	·	'''	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
		[[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
ı	ł	l	} ;	peptide	sequence	/=possible nucleotide deletion, \=possible
1]					nucleotide insertion
<u> </u>	<u> </u>	ļ		sequence		
					l	EMTNLKDIGLYNLRNITRG\AIRIEKNADLCYL
)]	1	Į.		1	STVDWSLILDAVSNNYIVGNKPPKECGDLCP
I			1		,	GTMEEKPMCEKTTINNEYNYRCWTTNRCQK
		1				MCPSTCGKRACTENNECCHPECLGSCSAPDN
1]]	1 .		J	DTACVACRHYYYAGVCVPACPPNTYRFEGW
	1					RCVDRDFCANILSAESSDSEGFVIHDGECMQE
			!		1	CPSGFIRNGSQSMYCIPCEGPCPKVCEEEKKT
1]			KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA
		,				SELENFMGLIEVVTGYVKIRHSHALVSLSFLK
		Ì	j		l '	NLRLILGEEQLEGNYSFYVLDNQNLQQLWD
			1	1	}	WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE
					,	VTGTKGRQSKGDINTRNNGERASCESDVLHF
						TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK
1				1	į .	EAPFKNVTEYDGQDACGSNSWNMVDVDLPP
1						NKDVEPGILLHGLKPWTQYAVYVKAVTLTM
						VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS
1			l i			NSSSQLIVKWNPPSLPNGNLSYYIVRWQRQP
						QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT
						ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE
1						AEYRKVFENFLHNSIFVPRPERKRRDVMQVA
}						NTTMSSRSRNTTAADTYNITDPEELETEYPFF
İ						ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE
1						KLGCSASNFVFARTMPAEGADDIPGPVTWEP
1 .						RPENSIFLKWPEPENPNGLILMYEIKYGSQVE
ŀ						DQRECVSRQEYRKYGGAKLNRLNPGNYTARI
ļ						QATSLSGNGSWTDPVFFYVQAKRYENFIHLII
-						ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN
· .						GVLYASVNPEYFSAADVYVPDEWEVAREKIT
						MSRELGQGSFGMVYEGVAKGVVKDEPETRV
1						AIKTVNEAASMRERIEFLNEASVMKEFNCHH
						VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR
		ĺ				SLRPEMENNPVLAPPSLSKMIQMAGEIADGM
					1	AYLNANKFVHRDLAARNCMVAEDFTVKIGD
						FGMTRDIYETDYYRKGGKGLLPVRWMSPESL
						KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ
						GLSNEQVLRFV\MEGGLLDKPDNCPDMLFEL
1						MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE
1					İ	VSFYYSEENKLPEPEELDLEPENMESVPLDPS
						ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF
1				i	}	DERQPYAHMNGGRKNERALPLPQSSTC
821	2171	Α	6691	106	825	GRVLFRGCGVGHKGOVLMGTFILAODWLSE
		·				SNHVFCVSSMLRLOKRLASSVLRCGKKKVW
		1		ĺ		LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT
				ļ		VHSRARCRKNTLARRKGRHMGIGKRKGTAN
						ARMPEKVTWMRRMRILRRLLRRYRES/KRYR
		- [ĺ	1	ſ	ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH
		1	ļ			IHKLKADKARKKLLADQAEARRSKTKEARK
	ļ	.]		J	J	RREERLQAKKEEUKTLSKEEETKK
822	2172	A	6715	772	21	DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI
			7,15			SGAKSSSS/RFTDSKRYEK\DFQ\SCFGLHETR\
		ļ	.]	j	}	SGDI/CNA/CVLL/LKRWKKLPAGSKK/NWNH
						VVDARAGPS/LKTTLKPKKVKTL/SGNRIK/ST
				İ		QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF
	.1	:				TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\
	1	- 1	ŀ	ļ		
		!	İ	ľ		TYWKRQKICCGNYKGRFGEVLIDTHLFKPCC
823	2173		6727		4063	SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
043	41/3	A	0/2/	3	4003	PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG
	. [1	1	[NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA
	ŀ	1				SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS
L			<u>l</u>			VGGQNPSTGGISADRTQGNIGCGGDTDPGQS
						· · · · · · · · · · · · · · · · · · ·

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
иепсе				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
		,				SSQPSQDGQESNVPSVGSLADPDYLNTPQMN TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL
	,					SDSVMNIFKDRNFDSCCICACNMNIKGADVG LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL FLEDELDIFGKNSDIGQAAERRLMMCQSTFL PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN FLDYISSNNRQTLPCVSWSYDRVQADNNDY
						WTECFNALEQGRQYVDNPTGGKVDEALVRS ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT
						ISPFSLPFWERLLLDPYGGHRDVAYTVVCPEN EALLEGAKTFFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWFNQPW SGEENDNHSRLKLYAQVCRHILAPYLATLQL
						DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV PPVSSSASAPGISQISTTSSSGFSGSVGGQNPST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVVIYM
						VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP AASIEMTLKNPERPPIQLYSPPFILAPIKDKQT
·						ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV
						TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP
924	2174		6722	2440	365	LALGYFVSTAKAENLPQWFWSSCPQAQMQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNALMNIL
824	2174	А	6732	2440	365	VEEGLGRRKTIPPGGRRGPVIPARPGFDSVRR RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD
						QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP
						NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYG\KLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI
						LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA
						KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD G RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARAITDYLQ ASAITRIPSYRYRYQRRSRSSSRSTEPSHSRDA
					'	SPYGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV
826	2176	A	6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSPQFAH LEFHLCHHILDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFBGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRILTL DNIFLCVAPAGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I-Isoleucine, K-Lysine, L-Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		Į.		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	_	/=possible nucleotide deletion, \=possible
]		sequence]	nucleotide insertion
						VGKGIMGVFTKPIGGAAELVSQTGYGILHGA
		i		1	ļ	GLSQLPKQRHQPSD\VHADQAPNSHVKYVW
1		1	1	ł		KMLQSLGRPEVHMALDVVLVRGSGQEHEGC
						LLLTSEVLFVVSVSEDTQQQAFPVTEIDCAQD
1		l]			SKQNNLLTVQLKQPRVACDVEVDGVRERLSE
	ĺ					QQYNRLVDYITKTSCHLAPSCSSMQIPCPVVA
ľ	ľ	ļ				AEPPPSTVKTYHYLVDPHFAQVFLSKFTMVK
						NKALRKGFP
827	2177	A	6748	2	1662	FVGAPRRGNPFGSPGNPGRHQGPCHRPRGTK
	:					ASGVSPTLWRPQAAATGLEMPSSGRALLDSP
1						LDSGSLTSLDSSVFCSEGEGEPLALGDCFTVN
		l				VGGSRFVLSQQALSCFPHTRLGKLAVVVASY
1			1			RRPGALAAVPSPLELCDDANPVDNEYFFDRS
					:	SQAFRYVLHYYRTGRLHVMEQLCALSFLQEI
		·				QYWGIDELSIDSCCRDRYFRRKELSETLDFKK
						DTEDQESQHESEQDFSQGPCPTVRQKLWNIL
1						EKPGSSTAARIFGVISIIFVGVSIINMALMSAEL
1						SWLDLQLLEILEYVCISWFTGEFVLRFLCVRD
1						RCRFLRKVPNIIDLLAILPFYITLLVESLSG\SQT
i 1						TQEL\ENVGAHCPGCLRLLRAL\RMLKAWGR
						HSTGLRSLGMTITQCYEEVGLLLLFLSVGISIF
						STVEYFAEQSIPDTTFTSVPCAWWWATTSMT
i						TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI
			1	İ		AIINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER
1					•	ASTRSSGGDDFWF
828	2178	A	6786	5672	1360	GTHPASSGPVPLPPAAVSAATREELGEPVPFV
			0.00	20,2		TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ
1						EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS
						ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK
1						NGAGNGAKGKGKRERSISADSFDORDPGTPN
						DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT
						APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS
						TEMANKAAEAVLKGOVETIVSFHIONISNNK
1 !			}			TERSTAPLNTQISALRNDPKPLPQQPPAPANQ
		1				DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE
			j	ļ		SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN
				ŀ	!	NRAVTPVSQGSNSSSADPKAPPPPPVSSGEPPT
	ļ				İ	LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP
					i	DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI
			1	j	į	QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR
						DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP
}					ļ	EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ
	1	1	' l		l	DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP
	1	,		,	İ	GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ
]]	Į	J	J	j		MRLPGFAGMINSEMEGPNVPNPASRPGLSGV
	}		l	ļ		SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP
			ļ	İ		NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR
	4		i			PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP
1 1)	1	ļ	ļ		LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM
		į				KGDVNLNVNMGSNSQMIPQKMREAGAGPEE
]	1	-				MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE
1	ľ	1	İ	ľ	1	YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ
1		i			1	RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG
	-	1		. [İ	LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ
	j]		SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS
. 1	ſ	- 1	ľ	ĺ	İ	PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG
1						
]		Ì]	1		
			Ì			GPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL SQNPLSIMMSR\MSKFAM\PS\SNPG\YNHDAI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	испсе	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ	}	}	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				•		KTVASSDDDSPPARSPNLPSMNNMPGMGINT
1	}	ł	}		ļ	QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ
	İ	ļ			•	MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH
		i				GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP
						HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQ
	{					SSADAALCKPGGPGGPDSFTVLGNSMPSVFT
1	j				,	DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ
1		l				YFPRGEVPGRKQPQGPGPGFSHMQGMMGEQ
1		ľ	1			APRMGLALPGMGGPGPVGTPDIPLGTAPSMP
i		l				GHNPMRPPAFLQQGMMGPHHRMMSPAQST
]		1		MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT
920	0120	<u> </u>	(505	422		HPGPVGSPGMMMSMQGMMGP\NRTS
829	2179	A	6797	433	3	ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH
						GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP
		l				ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG
		1				AAASSTSSGASCPPVPASSRWGVRSRTRSGSG
830	2180	Α	6800	3	1911	GEREPRDRPSERPRLV
830	2100	^	0600	3 ,	1911	LPERAFGPRTPRAPRRRRRLLLSPPPRPPPPL
			ļ		•	DREPRAPGPWLCPSRAGTAQDPARIRERRGR VAGGAAGPAMELRARGWWLLCAAAALVAC
		ļ				ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPO
1		1				AEISGEHLRICPOGYTCCTSEMEENLANRSHA
		ĺ				ELETALROSSRVLQAMLATQLRSFDDHFQHL
						LNDSERTLQATFPGAFGELYTQNARAFRDLY
						SELRLYYRGANLHLEETLAEFWARLLERLFK
1 1						QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\
-						RELRLRAT\RA\FVAAR\SFVQGLGVAS\DVVR
1 1				'		KVAQVPLG\PEC\SRAVIEAGSYC/ALHCVGVP
						GARPCPDYCRNVLKGCLANQADLDAEWRNL
						LDSMVLITDKFWGTSGVESVIGSVHTWLAEA
						INALQDNRDTLTAKVIQGCGNPKVNPQGPGP
						EEKRRRGKLAPRERPPSGTLEKLVSEAKAQL
ľ						RDVQDFWISLPGTLCSEKMALSTASDDRCWN
						GMARGRYLPEVMGDGLANQINNPEVEVDIT
						KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF
l						QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS
						SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL
<u> </u>						LPLLLFLALTVARPRWR
831	2181	Α	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA
						EGRLREKLFSGYDSSVRPAREVGDRVRVSVG
						LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS
				1	·	WDPAEHDGIDSLRITAESVWLPDVVLLNNND
						GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS
						IQVTYFFFDWQNCTMVFSSYSYDSSEVSLQT
						GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS
				[RLIQPPGDPRGGREGQRQEVIFYLIIRRKPLFY
1 1	i			ł	i	LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF ALLTLTVFLLLLADKVPETSLSVPIIIKYLMFT
	ļ		Ì			MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV
				ļ		RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP
]			-		ļ	GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL
		1	1	ļ		SAPDLRRFIDGPNRAVALLPELREVVSSISYIA
	ł	1	1	-[ROLOEQEDHDALKEDWOFVAMVVDRLFLW
						TFIIFTSVGTL\VIFLDATYHLPPPDPFP
832	2182	A	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK
		**	JU2-T		.3//	ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY
			ļ	Ì		DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN
	ļ	1	1		}	GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI
				ļ	•	KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAE
						KPIENRDFLKNSKILEICDNVTMYWINPTLIIS
	1					

SEQ ID NO: of nucl-	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, E=Phenylalanina G=Glycina H=Victidia
eotide	peptide seq-	1	USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence	l		914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan.
}	[residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
		·		peptide	- Sequence	/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion
					1	GTFAKQLHHNFAPIILVSELQDFEEEGEDLHFP
}		j		ļ	ļ	ANEKKGIEQNEQWVVPQVKVEKTRHARQAS EEELPINDYTENGIEFDPMLDERGYCCIYCRR
į		1				GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV
						IMPCNWWVARMLGRV
833	2183	A	6846	116	602	EAEGEQVCGAKCCGDAPHVENREEETARIGP
						GVMESKEERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRRFRVR
						QPILQYRWDIMHRLGEPQARMREENMERIGE
	1	ĺ				EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH
924	2104	ļ., —	(05)			DEFC\LMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSREESPAPSRAPASASLWRRLVVVEAKMAA
			[HAAAAAQAAAAQAAHAEAADSWYLALLGF
ĺ					,	AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA
1		ŀ				RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS
ł		ŀ				QQIPQFEDVKFEAASLLSELYCQENSVDAAKP
						LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK
ļ		į				GMLLLMERKLQEVHPLLTLCGQIVENWQGN
		Ì				PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC
						LKQLQQCIQTISTLHDDEILPSNPADLFHWLP
						KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMLDCSPILSSFQVILLEHIIM
					٠	CRLVTGHKATALQEISQVCQLCQQSPRLFSN
						HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL
						RLTNHQELWAFIVTNLASVYIREGNRHQEVV\
						LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA
						CSLVLLGHIFYVLGNHRESNNMVVPAMQLAS
						KIPDMSVQLWSSALLRDLNKACGNAMDAHE
						AAQMHQNFSQQLLQDHIEACSLPEHNLITWT
835	2185	Α	6855	334	1268	DGPPPVQFQAQNGPNTSLASLL PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL
	2105		0033	33.	1200	SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS
						AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE
1						LELVTSAKALNDTQKLACLIGVEGGHSLDNS
		-				LSILRTFYMLGVRYLTLTHTCNTPWAESSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV
						DLSHVSDAVARRALEVSQAPVIFSHSAARGV
						CNSARNVPDDILQLLEEERWAFVMVSLFHGE
						LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI
836	2186	Α	6862	315	11	GGDYDGAGKYRKKTTCKAPWRTSSRMSS PPRSRPSCWRKKVGPGRPWWWGGTGPPGOG
					- -	RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL
	J	•				PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT
027	2107		6962		1615	GPGSGRPPP
837	2187	A	6863	2	1615	VLRGQRGPAGGLÄEERRRGRNEWRIHDVIT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC
						NEDGLTALHQCCIDNFEEIVKLLLSHGANVN
			İ	ļ	İ	AKDNELWTPLHAAATCGHINLVKILVQYGA
			.			DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY
[ľ		ľ	QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH
						GVRVDVKDWDGWEPLHAAAFWGQMQMAE
]					ļ	LLVSHGANLNARTSMDEMPIDLCEEEEFKVL
	}	.	1		. }	LLELKYHKHDVIMKSQLRHKSSLSRRTSHRQA
	1					S/SVGKVVRRTQPVGTGPNL\YRKEYE/GEEAI
	اــِــــــــــــــــــــــــــــــــــ					LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
		hod				
NO: of	NO: of	пов	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	•		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	}		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			1	peptide		/=possible nucleotide deletion. \=possible
				sequence		nucleotide insertion
			 			PNPRLEK\PVLLSEFPTKIPRGELDMPVENGLR
]	1			
1						APVSAYQYALANGDVWKVHEVPDYSMAYG
1	ļ]			NPGVADATPPWSSYKEQSPQTLLELKRQRAA
						AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI
1						GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI
						EEMEEKVHGCCRIS
838	2188	A	6865	6291	739	AGPLEPRVQGAMALQLWALTLLGLLGAGAS
	l i					LRPRKLDFFRSEKELNHLAVDEASGVVYLGA
1						VNALYQLDAKLQLEQQVATGPVLDNKKCTP
						PIEASQCHEAEMTDNVNQLLLVDPPRKRLVE
						CGQLLKGI\CALRALSNISLRLFYEDGSGEKSF
1	1				'	VASNDEGVATVGLVSSTGPGGDRVLFVGKG
		l .			}	NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT
1					,	
						YKAGYLSTNTQQFVAAFEDGPYVFFVFNQQD
				· •		KHPARNRTLLARMCREDPNYYSYLEMDLQC
						RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF
					'	SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN
1					l i	ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK
					·	SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN
						LTAVTVAAENNHTVAFLGTSDGRILKVYLTP
						DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY
1						AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY
						CGWCVVEGRCTRKAECPRAEEASHWLWSRS
1						KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA
						LSEEDELLCLFGESPPHPARVEGEAVICNSPSS
					-	IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF
		-		,		
1 '				i		YDCRQAMSLEENLPCISCVSNRWTCQWDLR
1						YHECREASPNPEDGIVRAHMEDSCPQFLGPSP
						LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD
						LLKFMEPVTMQESGTFAFRTPKLSHDANETL
1						PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC
]						SLCRAANPDYRCAWCGGQSRCVYEALCNTT
				ĺ		SECPPPVITRIQPETGPLGGGIRITILGSNLGVQ
[]				i		AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA
				l		AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP
				J		KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED
						VRVTLNGVPCKVTKFGAQLQCVTGPQATRG
						QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE
						PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP
]						
.]	LQSWQPPREAESLQPMTVVGTDYVFHNDTK
]					1	VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT
						EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR
			ļ	ļ	j	GTNLNKAMTLQEAEAFVGAERCTMKTLTET
		1		•	ļ	DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF
					ļ	GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM
					ļ	VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG
, 1					ļ	LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG
	1				1	IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL
				1		DIPEPRRPVVEQALYOFSNLLNSKSFLINFIHT
				j		L'ENQPEFSARAKVYFASLLTVALHGKLEYYT
		- 1		ļ	į	
]		1				DIMHTLFLELLEQYVVAKNPKLMLRRSETVV
]	[J	J	J		ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI
						KHQVEKGPVDAVQKKAKYTLNDTGLLGDD
		Į	i			VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ
		ļ	1	1		VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG
		ŀ		1	1	STAQILSDLDLTSQREGRWKRVNTLMHYNVR
		Į		I	1	DGATLILSKVGVSQOPEDSQODLPGERHALL
				. !	ļ	EEENRVWHLVRPTDEVDEGKSKRGSVKEKE
]	. 1		l	*	RTKAITEIYLTRLLSVKGTLOOFVDNFFQSVL
	Į	ŀ		l		APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI
<u> </u>						A CHATTATE TELEPTENE CAERAINIQUED []

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	•	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	dance	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l donor	1	ļ	1 714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptonhan.
1	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}		1]	peptide	5-4	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
				504555		HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV
						DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL
1	1			1		YAKEISTYKKMVEDYYKGIROMVOVSDODM
						NTHLAEISRAHTDSLNTLVALHQLYQYTQKY
						YDEINALEEDPAAQKMQLAFRLQQIAAALE
1						NKVTDL
839	2189	Α	6872	1	1485	RARRLALQCHVCVCALTPGEQSGRRLPGOT
						WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH
						EDQTDCSSLRDENNKENYPDAGALVEEHAPP
ł			!			SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP
l						KSIFKAESGRSHGESQETEHVVSSQSECQVRA
l						GTPAHESPONNAFKCOET\VRL\OPRIDORTAT
j			j :			SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS
1						TQSVLA\DGTDSADPSPVHKDGQNEADSAPE
ļ						DLHSVGTSRLLL/YHITDGDNPTAVRHGCSL/F
1						SGQSQRFNLDPESAPSPPSTQQFMMPRSSSRC
						SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE
1 .						QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK
i '						QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP
Į						RENGKPEAAGPEPSSSGEETPDAALTCLKERR
[EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL
						STPSLIPTIVSQDTCMLLLCTDV
840	2190	A	6873	2	2054	FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL
						ENNRRSAACKRSPGTGDFSRNSNASNKSVDY
J						SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN
						YLKQPVVKEKEKKKYNVSKISQSKGQKEISV
						EKKHTWNASLFNSQIHMIAQRRDAMAHRILS
						ARLHKIKGLKNELADMHHKLEAILTENQFLK
						QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV
						KNLRQLLRKSQEKERTLSRKLRETDSQLLKT
						KDILQALQKLSEDKNLAEREELTHKLSIITTK
J						MDANDKKIQSLEKQLRLNCRAFSRQLAIETR
						KTLAAQTATKTLQVEVKHLQQKLKEKDREL
						EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD
						RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG
ł	1					NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY
						EDLSGEEKHLEVQILLENTGRQKDKKEDQEK
						KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR
						EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ
	j					RRHYSFTEATENLHHGLPASGGPANAGNMR
	ł			l		YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS
		ĺ		ļ		SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD
	,	Ì				QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA
841	2101	<u> </u>	6074		2067	FGDSKVTVVNSIKPSSPTEGKRKIII
041	2191	A	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG
			1		Į.	NAPAPGTPAASGWQPPTYHSGRAFSARYPRP
	İ				1	SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA
1		ł	ł		ł	DHAVRPLHGARGGQPPVPQQHVLERQVQLS
				ļ		QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE
	ŀ					DTPWSDQRPREGEGEPPRGQLQPSRPTRARG
	ļ	J	J	J		TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP
	1					REPRRTVSESVIAVKASFPSSALPPRTGVALG
				1		RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP
		1	. 1			SGSVGGPARPASGPRQAREASLVVTCRTNKF
	1		1	1	ł	RKNNYKWVAASSKSPRVARRALSPRVAAEN
					1	VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWOSEAG
			i		.	SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS
		- 1	- 1	İ		GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG
		ļ	l			TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN
						TOT PATTACOTTO TOTACO T

NO of NO of NO of No of bod Do NO: bod Do NO: bod bodation bodation bodation such bodation such bodation such bodation such	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decidide seq- USSN 09495 19496					beginning		D=Aspartic Acid, E=Glutamic Acid,
Sociation Soci				1.			
uence 094946 gortspondi gl to first anim oaid residue of peptide residue of peptide sequence Q-Glutamine, R-Arginine, S-Serine, S-Serine, Q-Glutamine, R-Arginine, S-Serine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, S-Serine, G-peptide sequence Q-Glutamine, R-Arginine, R-Serine, , R-Serine, G-peptide sequence Q-Glutamine, R-Arginine, R-Serine, G-peptide, R-R-R-R-R-R-R-R-R-R-R-R-R-R-R-R-R-R-R-				USSN	location	corresponding	
Serior S				r i	correspondi		M=Methionine, N=Asparagine, P=Proline,
amino acid residue of sequence				914		acid residue	
residue of peptide sequence		1		1		of pentide	
### ### ### ### ### ### ### ### ### ##	1						
	j	J		j		50425.00	
KGLVQVYKHRLCRIPPSRAHLPTKEASSLHA VRIAPTSKVIKTRYRLYKKTYASPIPSLS LISWRARRISISRILVIARIRPYASOGOKAQ POSPWRSKGYRCIGOVILYKVSANKLSKTSG QPSDAGSRPILIRTORLDPAGSCSRSLASRAVQ RSLAIRGARGREEKRYCMYYNRYGRONR GERCPVHIDPSKVAVCTRPVRGTOKKTDOTOC PPSHHVSKKAMPVCSYLRIGICOSNOCTYSHV YVSRKABVCSDPIKGYVPLIGGACKKKKTDOTOC PPSHHVSKKAMPVCSYLRIGICOSNOCTYSHV YVSRKABVCSDPIKGYVPLIGGACKKKKTDLOTOC PPSHHVSKKAMPVCSYLRIGICOSNOCTYSHV YVSRKABVCSDPIKGYVPLIGGACKKKKTDLOT CPSHARGACCFRGAQCQLLIERTQRRHSRRAAT SPARGYSDATARRYVSSHOPRKPSAAQPTX GTPSAAATAAAVAAPPHCPGGSASPSSKAG SSSSSSSSPASIDHEPSLGEALAAACSNR LCKLPSFISLQSSSPGAQPRVRAPRAFLTKDS SSSSSSSSPASIDHEPSLGEALAAACSNR LCKLPSFISLQSSSPGAQPRVRAPRAFLTKDS GOPHIKRRL 842 2192 A 6898 506 2071 WPDLVHTWSSEAMGSCCSCPDKDTVPDNH RKKKVNNDDDGBELGSGIMELTDTELLYT RKRDSVKWHYLLCRRYGYDSNLFSRESGERC GTGQGIFAKCARAELIPMLQEMQNSINI VVEEPVVERNNNQTELELYPRTPTTTTGFAA QLRNDVYRYPSGADSSHSSSRHSVSVSGARL PSVGESSTIPILJVAEEQVHTVYNTTGVQEER KKRTSVHVLLARRYSAGSSTFKEPSSIEDR DPQILLEPEGVKPVLGPTPVQKQLMEKEKLE QLGRDQVSGGANTHVEDHTGYDSDERRDAP SVNKLVYENNGSSIPSSRHSVVGSARL PSVGESSTIPILJVAEEQVHTVYNTTGVQEERRADP SVNKLVYENNGSIPSSRHSVGSRRDTTTYTYPTVPDIR RRSLEHRQLNVTQVDLEGGSSDNPQTFKTPT TRIAPTYPTASAHKUSTRENETDTYPVFDIBERRADP SVNKLVYENNGSSPSSRHSVARDALISTSTSTS TONINNSAQRATALLTYPENLPSLPFVFWFDIR RRSLEHRQLNVTQVDLEGGSSDNPQTFKTPT TRIAPTYPTASAHKAVGSRLEFVLQIPPVSRA VRTACVLTTAAVQLELTTPTQLYFNPDIR RRSLEHRQLNVTQVDLEGGSSDNPQTFKTPT TRIAPTYPTRTETTYPTRTPTTRYPTP TRIAPTYPTRTETTYPTNFDIR RRSLEHRQLNVTQVDLEGGSSDNPQTFKTPT TRIAPTYPTRTETTYPTTYPTSPT SQLPGPLAAAGAGFREFSLLTGGEPGVC WTHIFFLGRCSSSTWWKNNSSENTIPTYSVKY LEGGSFRGRTADPVFNFFFGFTPYPTGGARAAGA LLFGRSSNFTSASQAQGTGVCHHAR LLFVFSVETGF1 HAGQAGLELLTSGDPPASAS QSAGTTGRSQHTRPGYEFIPYPSAQQEDALKA LLFVFSVETGF1 HAGQAGLELLTSGDPPASAS QSAGTTGRSQHTRPGYEFIPYPTGGRSCVTHD TRAKKLRRIWRILEEKSVAGAVQTLLLRSQE GGVTSAAASTLEPPRPTUGRESCYCPUDDQ FRSSRSECAEVGWNLTYSRAGVSWVQAV PEDMTLIFRKKCKMECOVPAETLYDVTHDE TYRKKWDSNVERSSPLAPHAWNSVK HHKYPPRKDLVRAVSKSQCLAPHAPRETETTG QCAGRALGAAAAPALGSHEPPGTVWNSSSCLAPHAPRETETTG QCAGRALGAAAAPALGSHEPPGTVWKASSPLEP							
WITAPTIKVIKITRYRIVKKTPASPISAPPFILS LPSWARRISLISRIJLVINARIPPYASGGKAQ PGSPWWRSKGYRCIGGVLYKYSANKLISKTISG OPSDAGSPPULTROELDPAGCSCRSLASRAVQ RSLAIRGARGREKREYCKTYYNRFGRONR GERCPYHIDPEKVAVCTRYBGTCKSKTOTC FPSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YYSKRABVCSDFLKGYCFLGAKCKKHTILLC PPFARROACFROAGCOLLIRTOKRHSKRAAT SPARGFSDATARSR VSASHCPRKPSASOKPTR (TPSSAATTAAAVAAPPHCYGGAS PSSSKAS SSSSSSSPSPASLDHEAPPHCYGGAS PSSSKAS SSSSSSSSPASLDHEAPPLOGGAS PSSSKAS SSSSSSSSPASLDHEAPPLOGGAS PSSSKAS SSSSSSSSPASLDHEAPPLOGGAS PSPARTLYRDS GKPLHIKPRI. 842 2192 A 6898 506 2071 WPDLYHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELLLYT RREDSVEWHYLLCRRYGYOSNLFSESGRC QTGQGIFAFKCARAEELFMMLQEIMQNNSIN VVERPVVERNNNGTELLEVPTRYBTTTTTGAA QNLPNGYPRYPSGDASSHPSSNGSARL PSVGESSTHPLLVAREGVHTVVNTTOVQEE KNRTSVEVPLARVSNAESSTPKEPPSSIEDR SYNKLVVENNGLSIPELGVPTTWGTTOVQEE KNRTSVEVPLARVSNAESSTPKEPPSSIEDR SYNKLVVENNGLSIPEAGVATRATIVOGAE SYNKLVVENNGLSPFTVQKGLMEKKLE QLGRDQVSGGANTTEWTTGYDSDEBRDAP SYNKLVVENNGLSPFTVQKGLMEKKLE QLGRDQVSGGANTTEWTTGYDSDEBRDAP SYNKLVVENNGLSPFTVQKGLMEKKLE QLGRDQVSGGANTTEWTTGYDSDEBRDAP SYNKLVVENNGLSPFTVQKGLMEKKLE QLGRDQVSGGANTTEWTTGYDSDEBRDAP SYNKLVVENNGLSPFTVQKGLMEKKLE QLGRDQVSGGANTTEWTTGYDSDEBRDAP SYNKLVVENNGLSPTTYVNGER KNRTSVERPFTQFTTRYTTGAA QRAFTGTTAGSCHARVGSKLEFTLQTPVFNFDIR RRSLEHRQLNYQVDLEGGSDSDNPOTPKTPT TPLPQTTRYTTELLYAVDLECGSDSDNPOTPKTPT TPLPQTTRYTTELLYAVDLECGSDSDNPOTPKTPT TPLPQTTRYTTELLYAVDLECGSDSDNPOTPKTPT TPLPQTTRYTTELLYAVDLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTPT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTT TPLPQTTRYTTCHLAVGLECGSDSDNPOTPKTT TPLPQTTRYTTCHLAVGLECGSDSDNPOT			-		sequence		
LESWRARLSLSRSLVINRIRPYASGGGKA PGSPWWRSKGYRGGGV,YKVSANKISKTAG QPSDAGSSPLLRTGGLOPAGSCSRSLASRAC QPSDAGSSPLLRTGGLOPAGSCSRSLASRAC QPSDAGSSPLLRTGGLOPAGSCSRSLASRAC REALTRQARQRERKEYCYMYVRGGRONG GERCPYHIPPEKVAVCTREVRGTCKKTTOGT PFSHHVSKEKMPVCSPFLKGYCPLGAKCKKKFTILLC PPARARGACTRGAGOOLIHRTOKHSRRAAT SPAPOPSDATARSRVSASHOPPKPPSASQRTR QTPSSAALTAAAVAAPPLCTGGSASPSSKAS SSSSSSSSPASLDHEAPSLOGAALAAACSKR LCKLPSTISLQSSPPGAQRV KAARAALTKDS A 6898 506 2071	ļ						,
PGSPWARSKGYRCIGGVY,KVSANKLSKTSQ QPSDAGSRPLIKTGLIPAGGSCSRLASRAVQ RSLAIRQARQREKREVCMYYNRFGERON GERCEYHIDPEXVAVCTREVRGFCTGKTTGGTC PFSHHVSKEMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCFJGAKKHTLLC PPARRGACPRGAQCQLLHRTQKRHSRAAT "SPACOFSDATARSR VSASHGPKRFSASQRYTT QTPSSAALTAAA VAAPPHCPGGSASSPSSSRA SSSSSSSFASLDIEEAPSLQEAALAAACSNS LCKLPSPISLQSSFSFAQRPVRAPRATIKND LCKLPSPISLQSSFSFAQRPVRAPRAPLTKDS GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNCFKVINVDDDCNELGSGIMELTDTELLY RREDSVKWHYLCHRRYGYDSNLFFESGGRC QTGQGIFARCARAEELFNMLQEIMQNNSINI VVEEPVVERNINTOFLEVPRITRYTTPGFAA QNLPNGYPRYPSGGASSHPSSRHFSVGSALL PSVGEESTHPLLVAEGQVHTVYVTTGVQEER KNRTSVHPILARAVSNAESSTFKEEPSSIGDR DPOILLEPBGVKFVLGFTVQKQCMEKEKLE QLGRDQVSGGANNTEWDTGVDSDERRDAD SVNKLVYENNGLSPSAGVRAGRLTSTSTSD TQNINNSAGRTALLNYENLPSLPPVWEARK LSRDEDDNLGFKTVLGFTVGVQCMEKEKLE QLGRDQVSGGANNTEWDTGVDSDERRDAD SVNKLVYENNGLSPSAGVRAGRLTSTSTSD TQNINNSAGRTALLNYENLPSLPPVWEARK LSRDEDDNLGFKTPSLNGYHNLOPHENTV NTENTYPASAHKIEYSRRDCTTPTVPHDPINTP RPSLEBGDNLGFKTPSLNGYHNLOPHENTV NTENTYPASAHKIEYSRRDCTTPTVPHDPINTP RPSLEBGUNLGVDLGGGSDSDNQTPKTT TPLPQTTTERTELVAVIDIERTAAMSNLQKAL PRODGTSNKTRHNSTDLLR 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSRLEYLQPPVSRA AGRPGTTHASGKMAYQSRLEYLQPPVSRA LEEGSFRGRTADFVFMTLFGGFLMTLRGLFYCV LEEGSFRGRTADFVFMTLFGGFLMTLRGLFYCV LEEGSFRGRTADFVFMTLFGGFLMTLRGLFYCV LEEGSFRGRTADFVFMTLFGGFLMTLRGLFYCV SQLPGPLSALGAHGHLVYGELNHCGFFCVC WTILIFFLGRCISOSSMCAGEEPIC-PHELLRP SQLPGPLSALGAHGHLVYGELNHCGFFCVC WTILIFFLGRCISOSSMCAGEEPIC-PHELLRP SQLPGPLSALGAHGHLVYGELNHCGFFCVC WTILIFFLGRCISOSSMCAGEEPIC-PHELLRP SQLPGPLSALGAHGHLVYGELNHCGFFCVC WTILIFFLGRCTSOSTWVNKNISSTTYPESYF HERELAASTEPQGFPLLLPPSAAGEDALKA LM LYPENLGESSPRRTQSTRTRALGLFT LPMELAASTEPQGFPLLCHESSVQVPDQAD FRSRFSCCAEVGWNLTYSEAGVSWVAQAD FRSRFSCCAEVGWNLTYSEAGVSWVAQAD FRSRFSCCAEVGWNLTYSEAGVSWVAQAD FRSRFSCCAEVGWNLTYSEAGVSWVAQAD FRSRFSCCAEVGWNLTYSEAGVSWVAVAGSREE RMGGAGGGGGSDDDTTSLYAEAPHFREETET PGAGRALGAAAAPALSPHIPPGTGTAWHRARP LPSLALSSELSVOHADSUENDIBASTAVAESREE RMGGAGGGGGSDDDTTSLYAEAPHFREETE]			
QPSDAGSRPLLRTGRLDPAGSCSRSLASRAC RSLAIRQARQRRSKREYCMYNPRGRONG GERCPYHIDPEKVAVCTREVRGTCKKTTDGT PFSHHVSKEKEMPCYSPTLKGISSINGCPYSHV YVSRKAGVCSDFLKGYCFJGAKCKKKHTLLC FPSHHVSKEKEMPCYSPTLKGISSINGCPYSHV YVSRKAGVCSDFLKGYCFJGAKCKKKHTLLC FPPARRGACTRGAQCQLLHRTQKRHSRRAAL SPACPGDATARSKYASHOPRKSPASQRPTR QTPSSAALTAAAVAPPHCPGGSASPSSSKA SSSSSSSSFSALDEBAPSLQEAALAACSNR LCKLPSTSLQSSF9GAQRVRAFRAPLTKDS GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVYDNH RNEFKVINVDDDGNELGSGIMELITIFLLLYT RRABDSVKWHYLCLRRYGYDSNLFSFESGRC QTGQGGFAFKCARAELFINNLGLIMONISIN VVEEPVVERNHOTELEVPRTPRTTTTCPAL QTGQGFAFKCARAELFINNLGLIMONISIN VVEEPVVERNHOTELEVPRTPRTTTTCPAL QUGROVSSGGANTEWDTGYDSDERADA QNLPHOLYSTYHPLEARSTSTSDT QTGCGFAFKCARAELFINNLGLIMONISIN VVEEPVVERNHOTELEVPRTPRTTTTCPAL QUGROVSSGGANTEWDTGYDSDERADA QNLPHOLYSTYHPLEARSTSTSDT QTGCGFAFKCARAELFINNLGLIMONISIN VVEEPVVERNHOTELEVPRTPRTTTTCPAL QUGROVSSGGANTEWDTGYDSDERADA QNLPHOLYSTHPLLAEGUNTYVTTGVOCHER KNRTSYHVPLEARSVNASSSTYPLEFPSSIEDR DPQILLEPEGVKFVLGFTPVQKQLMEKEKLE QLGRQVSSGGANTEWDTGYDSDERADA QNLPHOLYSTHPLAEGUNTYVTTGVOCHER KNRTSYHVPLEARSVNASSSTYPLEFPSSIEDR DPQILLEPEGVKFVLGFTPVVQKQLMEKEKLE QLGRQVSSGGANTEWDTGYDSDERADA VVEEPVVERNHOTELSTPTVQKQLMEKEKLE QLGRQVSGGANTEWDTGYDSDERADA VYELFTCHTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1			l			
RSLAIRQARQREKREVGMYYNRFIGRONG GERCPYHIDPEKVAVCTRFVRGTCKSTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKHTILC PFSHAVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKHTILC PFPARRGACRGAQCQLLHRTQKRHSRAAAT SPAPGFSDATARSR VSASHGPRKPSASQRTT QTPSSSAALTAAAVAAPPHCGGSASSPSSSKAS SSSSSSSSPSALDHEAPSLQEAALAAACSN LCKLPSFISLQSSFSFGAQRVRAPRAPLTKDR LCKLPSFISLQSSFFGAQRVRAPRAPLTKDR GKPLHIKPRI 842 2192 A 6898 506 2071 WPDLVHTWSSEAMGSCCSCPDKDTVPDNH RNSFKVINVODDGNELGSGIMELTDTELIL YT RRDSVKWHYLCRRYGYDSNLFSSGGRC QTGQGFARCARAEELFNMLQEMQNNSIN VVEEPVVERNNHQPLELYPRTPRTTTTGFAA QNLPNGYPRYPSFODASSHFSSRHPSVGSARL PSVGEESTIPLLVAEEQVHTYVNTTGVQEE KNRTSVHYPLEARVSNASSSTKEEPSSIEDR DPQILLEPFGVKFVLGPTFVQKQLMEKEKLE QLGRDQVSGSGANTTEWDTGYDSDERRDAP SVNKLVYENINGLSPSASGVRGRUTTSTTD TQNINNSAQRRTALLNYENLPSLPPYWEAKK LSRDEDDNLGPKTPSLNGYHNNLDPMHTNY NTENNYTYPASAHKEYSRRBCDTFVFNFDID RRSLEHRQLNYIQVDLGGGSDSDNPQTPKTPT TPLPCFTRRTELLYAVDIBERTAAMSNLQKAL PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKNAYQSSLTEVQPPVSRA UVFLOPELYNNIGSSMCGAEEPLCPHELIR; RPSLEHRQLNYIQVDLELITPGQLYPPVSRA VTTACVLTTAAVQLELITPGQLYPPVSRA LEGSFRGRTADFYMFLFGGFMLLICGLFVS LVYLGPGLYNNIGSSMCGAEEPLCPHELIR; SQLPCPLSALGAHGFTLVGGHENHGFPGYCS WTHIFFLGRCISGSTWYNKNISENTTYFESYF SQLPCPLSALGAHGFTLVGGHENHGFPGYCS WTHIFFLGRCISGSTWYNKNISENTTYFESYF SQLPCPLSALGAHGFTLVGGHENHGFPGYCS WTHIFFLGRCISGSTWYNKNISENTTYFESYF SQLPCPLSALGAHGFTLVGGLHDHGOPGYCS WTHIFFLGRCISGSTWYNKNISENTTYFESYF SQLPCPLSALGAHGFTLVGGLHDHGOPGYCS WTHIFFLGRCISGSTWYNKNISENTTYFESYF TRAKKLRRURLEEKESVAGAVQTILLCRGLPGY LPSLASASTEPRETGESRTTRALGLPT LPMEKLAASTEPQGFRPYLCRESCVQVPDQD FRSRSECCEAEVGWITTYSRAQVSWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHOUP FRSRSECCEAEVGWITTYSRAQVSWVVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHOUP FRSRSECCEAEVGWITTYSRAQVSWVVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHOUP FRSRSECCEAEVGWITTYSRAQVSWVVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHOUP FRSRSECCEAEVGWITTYSRAQVSWVVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHOUP FRSRSECCEAEVGWITTYSRAQVSWVVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHOUP FRSRSECCEAEVGWITTYSRAQVSWVVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHOUP FRSRSECCEAEVGWALTSTRAGTGTWHKARE	ĺ	ſ		1			l
GERCPYHIDPÉKVAVCTREVRGTCKKTOGTC PFSHHYSEKNAVCSSYLEGICSNSNOPYSHY YVSRKAEVCSDFLKGYCHOAKCKKHTLLC PPDFARRGACPGOLHENGRHSRAAT SPAPGPSDATARSRVSASHGPEKPSASQRTR QTPSSAALTAAAVAAPPHCPGGSASSKAS SSSSSSSPALDHEAPSLQEAALAAACSNR LCKLPSFISLQSSYSRAQPRVRAPRAPLTKDS GKPLHIKPRL 842 2192 A 6898 306 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGRELTDTELLLYT RKRDSVKWHYLCLRRYGYDSNI-SFSESGRC QTGQGFAFKCARAELFMALGEMQNITYPNTTGVQER KNRTSVHVPLEARVSNAASSTREYGSARL QNLPNGYPRYSFGDASSHSPFSSGRC QTGQGFAFKCARAELFMALGEMQNITYPNTTGVQER KNRTSVHVPLEARVSNAASSTREYGSBARL QNLPNGYPRYSFGDASSHSFFSSGRC QLGRDQVSGSGANTIEWDTGNDSERRDAP SVNKLVYENINGLSIPSASGVRGRLTSTSTSD TQNINNSAQRATALLAYEALFSLPPVWEARK LSRDEDDALGRKTPSINGVHNLDPMHNYV NITENTYTVASAAHKLEYSRRDCTFTYPHOR RPSLEHRQLNVIQVDLEGGSDDNOTPRTPTT TPLPQTPTRTELLYAVIDIERTAAMSNLQKAL PRDDGTSRKTRINSTDLLL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPVSRA YTTACVLTTAAVQLELTTPGJVYNDELIFKHF QWRLITHFLFGGVGFNFLFNMIFLYSCM LEGGSFRGRTADVFMFLEGGFLMTLGCFPGVS LVFLGGLVNNGSSMGGABEPLCPHELLRP QWRLITHFLFGGVGFNFLFNMIFLYSCM LEGGSFRGRTADVFMFLEGGFLMTLGCFPGVS LVFLGGLVNNGSSMGGABEPLCPHELLRP SQLPGLSALGAHGFLVUGFLANGFSVGVTU AHCMELPGSSNSPASAQVAGITGVCHHAR LIFVFSVETGFLHACQAGLELTSGDFPASAS QSAGTGKSGPIRRYGTEPTYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENIGESLYPPWDGGRPCCVEMS TRAKKLRRWRILEKERSVAGAVOTILLRSQE GGVTSAAASTLSEPRRTGESSTRTRALGLFT LPMEKLAASTEPQGFRVLGRESVGVPDQDQ FRSFRSECEACVOWLTYSRAGVSWVVQAV EMBRILHKIKCRMECCOVPAETLYDVLIDIE YRKKWDSNVUETFOLARL TWONGYTSWR CPEPLKNRDVITLRSWLPAGADVIBMYSVK HPKYPPRKDLVRAVNSSQFLAFRAMSK MYKACLKYPEWGCKHLIPHEKPWLHTPEGSF LPSLALSSELSVQAADSLELBESAAVASSREE RMGGAGGGGSDDDTSLYABAPIERFRETETG PGAGRALGAAAPALSSLEIPDESAAVASSREE RMGGAGGGGSDDDTSLYABAPIERFRETETG PGAGRALGAAAPALSSLEIPDESAAVASSREE RMGGAGGGGSDDDTSLYABAPIERFRETETG PGAGRALGAAAPALSSLEIPDESAAVASSREE RMGGAGGGGSDDDTSLYABAPIERFRETETG PGAGRALGAAAPALSLEIPDFTWWIRRABP							
PFSHHVSKEKMPVCSYPLKGICSNSNCPYSHV YVSRKAPCCSDPLKGYCPLGKOKKHTILC PDPARRGACFRGACCQLLHRTQKRHSRRAAT SPAPGPSDTAKARSVASAHGPKSASQRFTR QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPASLDHEAPSLQEAALAAACSNR LCKLPSFIGLGSSPSRGAGPRAFAPLTKDS GKPLHKRRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNFKVINVDDDGNELGSGMELTDTELL LYT RKRDSVKWHVLCLRRYGYDSNL FSFESGRAC QTGGGIFAFKCARAEELFMILQEIMQNNSIN VVERPVVERNNHOTELEVPRTTTTPGFAA QNLPNGYPRYPSFODASSHPSSRFSRHFSVGSARL PSVGEESTHPLLVAEDQWITVTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPITPVQKQLMEKEKLE QLGRQVSSOSGANITEWDTGTSTSTSD TQNINNSAQRATALLAYYELFLPYWEAKK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENNYLVPASAHKEYSRRACHTSTSTSD TQNINNSAQRATALLAYYELFLPYWEAKK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENNYLVPASAHKEYSRRACHTSVHYDFUFTYT TPLPQTPTRRTELYAVDIDERTFYTYNTPOIR RPSLEHRQLNYIQVDLEGGSDSDNPOTPKTPT TPLPQTPTRRTELYAVDIDERTFYNTPOIR RPSLEHRQLANVIQVDLEGGSDSDNPOTPKTPT TPLPQTPTRRTELYAVDIDERTANSNLQKAL PRDDGTSRKATRINSTOLDL AGRGTTHASGKMAYQSIRLEYLQPPVSRA YTTACVLTTAAVQLELTPPQLYFNMELTKYCM LEEGSTRGRT ADVFNFLTGGFGMTLEGFPVS LVFLGGLYNNVGSSMGAEEPLCPHELLRP QWRLITNFLFFGFVGFNFLFNMELTKYCM LEEGSTRGRT ADVFNFLTGGFGMTLGFFFCSV LVFLGGLYNNVGSSMGAEEPLCPHELLRP QURALTHAGGAGALELLFOLECNGVIL AHCNLRJGSSNSPASASQVAGTTGVCHHAR LIFVFSVETGFLHAGQAGALELLFOLECNGVIL AHCNLRJGSSNSPASASQVAGTTGVCHHAR LIFVFSVETGFLHAGQAGALELLFOLECNGVIL AHCNLRJGSSNSPASASQVAGTTGVCHHAR LIFVFSVETGFLHAGQAGALELLFOLECNGVIL AHCNLRJGSSNSPASASQVAGTTGVCHHAR LIFVFSVETGFLHAGQAGALELLFOLECNGVIL AHCNLRJGSSNSPASASQVAGTTGVCHHAR LIFVFSVETGFLHAGQAGALELLFOLECNGVIL LYFRSECEAEVGWNLTYSRAGVSVWVQAV PMSTLHKKCRMECCDVPAETLYDVLHDIE YRKKLMSNVUEGESVQVPDQDQ FRSFRSECEAEVGWNLTYSRAGVSVWVQAV PMSTLHKKCRMECCDVPAETLYDVLHDIE YRKKLMSNVUEGTFOLATLYDVLHDIE YRKKLMSNVUEGTFOLATLYDVLHDIE YRKKLMSNVUEGTFOLATLYDVLHDIE YRKKLMSNVUEGTFOLATLYDVLHDIE YRKKLMSNVUEGTFOLATLYDVLHDIE YRKKLMSNVUEGTFOLATLYDVLHDIE YRKKLMSNVUEGTFOLATLYDVLHDIE YRKRLMSROUGHSSVQVQAV PMSTLHKKLEKESVAGAQVOILLLRSQE GGVTSAAASTLESPRRTGESTRITRALGLPT LPHALASBLEDYGAADALSBTEBERAVAESREE RMGGAGGGSDDDTSLYABAPKRETETT PGAGRALGAAAAALSLEHPERFETTET PGAGRALGAAAAAA	i						
### PAPERS ALTER AND ALTER	1	l					
### PIPARRGACPRGACQCULHRTOKRHSRRAAT SPAPOPSDATARSRVASAHGPRSAGQRPTE QTPSSAALTAAAVAAPHCPGGSASPSSSKAS SSSSSSSPASLDHEAPSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVAPRAPLTKDS #### LCKLPSFISLQSSPSPGAQPRVAPRAPLTKDS GKPLHIKPRL ### A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNFKVNIVDDDGNELGSGIMELTDTELLLYT RKRDSVKWHVLCLRRYGYDSNLPSFESGRAC QTGQGIFAFKCARAEELFMHLQEIMQNNSIN VVERPVVERNHOTELEVPRTTTPGPAA QNLPNGYPRYSPGDASSIPSSRIPSVGSARL PSVGEESTHPLLVAEQWHYNTGVOERE RNRTSVHYPLAAVSDASSIPSKEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKELLE QLGRQCYSGSGANNTEWDTGDSBERKALE RNRTSVHYPLAARVINGLSIPSASGVRGGLTSTSTSD TQNRNNSAQRTALLATYPELIPSPWEAKK LSRDEDDNLGPKTPSLNGYHNNLDPMENNYV NTENNYTVPASAHKIEYSRRAFTALALLYPELIPSPWEAKK LSRDEDDNLGPKTPSLNGYHNNLDPMENNYV NTENNYTVPASAHKIEYSRRAFTALALLYPELIPTVWFDRF RPSLEHRQLNYIQVDLEGGSDSDNPQPTRTPT TPLPQTPTRRTELYAVDIBETRAMSNLQKAL PRODGTSRKITRINSTOLPL ##################################		l					l
SPAPGPSDATARSKYSASHGPRKPSSQRPTR	-	Ì	ľ	1			YVSRKAEVCSDFLKGYCPLGAKCKKKHTLLC
GTPSSAALTAAAVAPPHCPGGSASPSSKAS SSSSSSSPSPSALDHEAPPLOEAALAAACSNR LCKLPSFISLQSSFSPGAQPRVRAPRAPLTKDS GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVNVDDDOENELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGSNLFSFSGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNHQTELEVTRTPTTPTOFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAAEQVHTYVNTTIGVGER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEFEGVKFVLGFTPVQKQLMKEKLE QLGRDQVSGGANNTEWDTGYDSDERTDAF SVNKLVVENNGLSIPSASCVRRGRTATTSTSD TQNINNSAGRTALLINYBLPSLPPVWEAK LSRDEDDNLGPKTPSLNGYHNNLDPMHNVV NTENVTVPASAHKIEYSRRADCTPTVFNFDIR RPSLEHRQLWIYQVDLEGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDLERTAAMSNLQKAL PRDDGTSRKKTHNSTUDL 843 2193 A 6919 2 663 AGRFGTHASGKMAYQSLRLEYLQIPPVSRA VTTACVLTTAAVQLELITYLQYPPVSRA VTTACVLTTAAVQLELITYLQYPPVSRA VTTACVLTTAAVQLELITYLGYPPVSRA LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS LVFLGFGGLYNNVGSMCGAEEPLCPHELLRP SQLPGFLSALGAHGFLVVGELNHCOPFGYCS WTHEFFLGRIGSGTVWNGSMCGAEEPLCPHELLRP SQLPGFLSALGAHGFLVVGELNHCOPFGYCS WTHEFFLGRIGSGTVWNGSMCGAEEPLCPHELLRP LIPVFSVETGFLHAGQAGLELTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 844 2194 A 6928 902 366 HRLCMPQGACGERMEFFSILLFGLECNGVIL AHCNLENPESSNPSALSAGVAGTTGVCHHAR LIFVFSVETGFLHAGQAGLELTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYFENLGESKESVAGAVQTLLLRSQE GGVTSAAASTLSEPRRTQESRTRTRALGLFT LPMEKLAASTEPQGFRPVLERSVQVPDOD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRITHKLKCRMECCDVPAETLVVLHDIE YRKKWDSNVHETFDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPKPLKNROVHTETDLARLTVNADVGYYSWR CPK	1						PDFARRGACPRGAQCQLLHRTQKRHSRRAAT
SSSSSSSSSPASLDHEAPSLQEAALAAACSNR LCKLPFSISLOSSPSPAQPRVRAPRAPITKDS GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDMH RKRSVKWHYLCLRYGYDSNLSFSGRRC QTRAKKKVNVDDDGNEGSGIMELTDTELLLYT RKRDSVKWHYLCLRRYGYDSNLSFSGRRC QTGQGIFAKCARAEELFIMILQEIMQNSIN VVEEPVVERNIHQTELEVPRTPRTPTTPGFAA QNLPNOYPRYPSFGDASSHPSSRHPSSRHSYGSRAL PSVGESSTHPLLVAEEQVHTVVNTTGVQEER KNRTSVHVPLEARVSNAESSTFKEEPSIEDR DPQILLEFEGVKFVLGFTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERDAP SVNKLVYENINGLSPSACYRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNLDPMENYV NTENVTVPASAHKIEYSRRRDCTFTVFNTDIR RPSLEHRQLNYIQVDLEGGSDSDNPOPTPKTPT TPLPQTPTRTELLYAVUDELTAMSNLQKAL PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPQLYFNPELIFKHP QIWKLITNFLFFGPVGFNILFNMIFLYRYCRM LEEGSSTGRATADFVTMELGPLNICGPFQYCS WTHIFFLGRCISQSTWWNKNSENTTYFESYF EVALUATION LVFLGPGLYNNIGSSMCGAEPLCPHELLRP SQLPGPLSALGAHGFILVAGGTMITGGLFVS LVFLGPGLYNNIGSSNGCAEEPLCPHELLRP SQLPGPLSALGAHGFILVAGGVAGTIGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGSSNSPASASQAGTGGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGSSNSPASASQAGTGGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESSNSPASASQAGTGGVCHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGKSQHTRPGYEFIIPYSAAQEDALKA LM 846 LYPENLGELSFPILLLPPPWDDGGRPCVEMS TRAKKLRRIWRLEEKESVAGAVQTLLLIRSQE GGVTSAAASTLSEPPRTQGSRTRTALGLFT LPMEKLAASTEPQGRPPUJGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWQAV EMDRITHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVETFDLARLTVNADVGYT'SWR CPKPLKNROVITLSWLPMGADTJUNNYSVK HPKYPPRKOLVASTEPQURFUVJGRESVQVPDQD FRSFRSECEAEVGWNLTYSRAGVSVWQAV EMDRITHKIKCRMECCDVPAETLYDVLHDIE YRKWDSNVETFDLARLTVNADVGYT'SWR CPKPLKNROVITLSWLPMGADTJUNNYSVK HPKYPPRKOLVASVGTGFVLJGRSTØKSCVIT YLAQVDPKGSLPKWVNKSSGFLARKAMK MYKACLKYPEWKQKHLYPFGWWHRAEP LPSLALSVELSVQHADSLENIDESAVAESREE RIMGGAGGEGGSDDDTISLYAEAPHRRETETO PGAGRALGAAAPAALSPFGTWWHTRAEP PGAGRALGAAAPAALSPFFGTWWHRAEP PGAGRALGAAAPAALSPFFGTWWHTRAEP PGAGRALGAAAPAALSPFFGTWWHRAEP PGAGRALGAAAPAALSPFFGTWWHRAEP PGAGRALGAAAPAALSPFF	Ì	į.					SPAPGPSDATARSRVSASHGPRKPSASQRPTR
LCKLPSFISLQSSPSGAQPRVRAPRAPLTKDS GKPLHIKPIL 842 2192 A 6898 506 2071 WPDLVHTWSSEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDOINELGSGIMELTDIELLLYT RKRDSVKWHYLCLRRYGYDISNLFSFESGRC QTGQGIFAFKCARAEELFINILQEIMQNISIN VVEEPVVERNHOTELEVPTRIPTTPTGFAA QNLPINGYPRYPSFGDASSHPSSRHPSVGSARI. PSVGESSTHPILVAEEQYHTVVNTTGVQEER KNRTSYNVPLEARVSNAESSTPKEEPSSIEDR DPQILLEFEGVKFVLGFTPVQKQLMEKEKLE. QLGRDQVSGGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSPSASGVRRGRITSTSTSD TQNINNSAGRTALLNYFULPSLPPVWEARK LSRDEDDINLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTFTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPGTPTRRTELVAVIDLEGGSDSDNPQTPKTPT TPLPGTPTRRTELVAVIDLEATAAMSNLQKAL PRDDGTSRKTRHNSTULPL 843 2193 A 6919 2 663 AGRPGTHASGKMAYQSLRLEYLQIPPVSRA LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS LVFLGPGLYNNIGSSMCAGEEPLCPHELLRP QIWRLITNELFGPVGFNLENMIFLXYCCM LEEGSFRGRTADFVFMFLGGFLMTLFGLFVS LVFLGPGLYNNIGSSMCAGEEPLCPHELLRP SQLPGLSALGAHGFLVVGELNHCOPPGYCS WTHIFFLGRISSTTWVNKNSENTITFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERMEFSILLFGLECNGVIL ALCHLERSSNSPASASQAGTGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGKSQHTRPGYEFIIPYSAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESNSPASASQAVGTTGUCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGKSQHTRPGYEFIIPYSAQEDALKA LM 846 2195 A 6939 1660 317 LYPENLGERESSVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLFT LPMEKLAASTEPOGPRPVGRESVQVPDODD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMBRITHKIKCRMECCDVPABILTVVLHDIE YRKKWDSNVEIFTDLARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYINNYSVK HPKYPRKOLVASVIGTGVCHINNYSVK HPKYPRKOLVAVSIGTGVCHINNYSVK HPKYPRKOLVAVSIGTGVCHINNYSVK HPKYPRKOLVAVSIGTGVLIJGSTOPKSCVIT YLAQVDPKGSLPKWVNKSSQFLAFKAMKK MYKACLKYPEWQKICHLJARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYINNYSVK HPKYPRKOLVAVSIGTGVLIJGSTOPKSCVIT YLAQVDPKGSLPKWVNKSSQFLAFKAMKK MYKACLKYPEWQKICHLJARLTVNADDVGYYSWR CPKPLKNRDVITLRSWLPMGADYINNYSVK HPKYPRKOLVAVSIGTGPVLIJGSTOPKSCVIT YLAQADPKGSLPKWVNKSSQFLAFKAMKK MYKACLKYPEWQKICHLJARLFWWLHPEDEG LPSLALSELSVQHADSLENIDESAVAESREE RIMGGAGGEGGSDDDTISLYAEAPHRRETETG PGAGRALGAAAPALSLEPFGTWWHTRAEP	ļ	l .		i i			QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS
	i				'		SSSSSSSPPASLDHE\APSLQEAALAAACSNR
	1	l	l	1 .			
RNKFKVINVDDDGNELGSGIMELTOTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEYPRTPTTTPGFAA QNLPNGYPRYPSFGDASSHFSSRHFSVGSARL PSVGESTHPLLVAEEQVHTYVNTTGYQEER KNRTSVHYPLEARVSNAESSTPKEEPSSIDGR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVINKLVYENINGLSIPSASGVRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSIPPVWEARK LSRDEDDALGPKTPSISPSASGVRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSIPPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSIPSPVVEARK LSRDEDDALGPKTPSIPSPVVEARK LSRDEDALGPKTPSIPSVPVGENTHATOKAL PRIDGTSRIKTRINSTDLPL LEGSTRGRTADFVPMTAVGIENTLGGPFVSTA AGRPGTTHASGKMAYGSLRLEYLQIPPVSRA YTTACVLTTAA VQLELITIPFQLYFNPELIFKHF QIWRLITINFLFFGPVGFNFLFNMIFLYRYCRM LEGSTRGRTADFVPMLFGGFLMTLFGLFVS LVFLGPGLYNN/GSSMCCAEEPLCPHELLRP SQLFGPLSALGAHGIFVGELNHCOPPGYCS WTHIFFLGRCISQSTWWNKNISENTYFESYF SQLFGPLSALGAHGIFUPSALGAGGEAGEALKA LIFVFSVETGFLHAGQAGLELTSGDPPASAS QSAGTGKSGYTTRYGSNSPASASQVAGTGVCHHAR LIFVFSVETGFLHAGQAGLELTSGDPPASAS QSAGTGKSGYTTRYGSNSPASASQVAGTGVCHHAR LIFVFSVETGFLHAGQAGLELTSGDPPASAS QSAGTGKSGYTTRYGTPYSAGGDAKKA LM 845 2195 A 6939 1660 317 LYFENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWBLEEKESVAGAVQTLLLRSQE GGVYTSAAASTLSEPPRRTOESRTTRALGLPT LPMELLASTEPQGFPVLGRESVQVPDDQD FRSSRSECEAEVGWNLTYSRAGVSVWVQAV BMDTLHKIKCRMECOPAETLYDVLRDIE YRKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIMNYSVK HPKYPPRRDLVRANSIOTYLIGSCSVTT YLAQVDPRGSLPFWLVNKSSOFLAFKAMKK MYKACLKYPEWKGMPFKPWLUPPEGSP LPSLALSELSVQHADSUENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHFRFETET PGAGRALGAAAAPLSPLHPPFDTWWHRARP	i		ļ				
RNKFKVINVDDDGNELGSGIMELTOTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEYPRTPTTTPGFAA QNLPNGYPRYPSFGDASSHFSSRHFSVGSARL PSVGESTHPLLVAEEQVHTYVNTTGYQEER KNRTSVHYPLEARVSNAESSTPKEEPSSIDGR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVINKLVYENINGLSIPSASGVRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSIPPVWEARK LSRDEDDALGPKTPSISPSASGVRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSIPPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSISPVWEARK LSRDEDDALGPKTPSIPSPVVEARK LSRDEDDALGPKTPSIPSPVVEARK LSRDEDALGPKTPSIPSVPVGENTHATOKAL PRIDGTSRIKTRINSTDLPL LEGSTRGRTADFVPMTAVGIENTLGGPFVSTA AGRPGTTHASGKMAYGSLRLEYLQIPPVSRA YTTACVLTTAA VQLELITIPFQLYFNPELIFKHF QIWRLITINFLFFGPVGFNFLFNMIFLYRYCRM LEGSTRGRTADFVPMLFGGFLMTLFGLFVS LVFLGPGLYNN/GSSMCCAEEPLCPHELLRP SQLFGPLSALGAHGIFVGELNHCOPPGYCS WTHIFFLGRCISQSTWWNKNISENTYFESYF SQLFGPLSALGAHGIFUPSALGAGGEAGEALKA LIFVFSVETGFLHAGQAGLELTSGDPPASAS QSAGTGKSGYTTRYGSNSPASASQVAGTGVCHHAR LIFVFSVETGFLHAGQAGLELTSGDPPASAS QSAGTGKSGYTTRYGSNSPASASQVAGTGVCHHAR LIFVFSVETGFLHAGQAGLELTSGDPPASAS QSAGTGKSGYTTRYGTPYSAGGDAKKA LM 845 2195 A 6939 1660 317 LYFENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWBLEEKESVAGAVQTLLLRSQE GGVYTSAAASTLSEPPRRTOESRTTRALGLPT LPMELLASTEPQGFPVLGRESVQVPDDQD FRSSRSECEAEVGWNLTYSRAGVSVWVQAV BMDTLHKIKCRMECOPAETLYDVLRDIE YRKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIMNYSVK HPKYPPRRDLVRANSIOTYLIGSCSVTT YLAQVDPRGSLPFWLVNKSSOFLAFKAMKK MYKACLKYPEWKGMPFKPWLUPPEGSP LPSLALSELSVQHADSUENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHFRFETET PGAGRALGAAAAPLSPLHPPFDTWWHRARP	842	2192	Α	6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH
RKRDSVKWHYLCLRRYGYDSNLFSFESGRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVENNHOTELLEVPRTRTTTTGFAA QNLPNGYPRYPSFGDASSHPSSRHFSVGSARL PSVGEESTHPLLVABEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTFKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDEERDAP SVNKLVYENINGLSPSASGVRGCRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMENYV NTENVTVPASAHKIEYSRRDCTFTVFNFDIR RSLEHRQLNYIQVDLEGGSDSDNPQTFKTPT TPLPQTFTRRTELYAVDIDERTAAMSNLQKAL PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQUPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS LVPLLGPGLYNNGSKGABEPLCPHELLRP SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHEFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPGQAGCERMEFSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGTTGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVOTLLIRSQE GGVTSAAASTLSEPPRTQESRTRTRALGLPT LPMEKLAASTEPGGRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECOVPAETLYDVLHDIE YYKKWDSNVIETBDLARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVNNSSQFLAPKAMKK MYKACLKYPEWKOKHLPHFEFWLUFDGSP LPSLALSSELSVQHADSLENIDESAVALESREE RWGGAGGGGSDDDTSLYAEAPHRFFETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	1 5.2		**	0070	500	20	l
QTGQGFAFKCARAEELFMMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPOFAA QNLPNGYPRYPSFGDASSHPSSRIPSVGSARI. PSVGEESTHPILVAAEQVHITYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDDERRADA SVNKLVYENNGLSFBASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAIKIEYSRRDCTFTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTFT TPLPQTPTRRTELYDUIBERTAMSNLQKAL PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLREYLQIPPVSRA YTTACVLTTAAVQLELITPPQLYPNPELIFKHF QIWRLITNFLFGGPVGFNFLFNMFLYRYCRM LEEGSFRGRTADFWHFLFGGFPGVCS MTHIFFLGRCISGSTWWNKNSSNTTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISGSTWWNKNSSNTYFESYF TRAKKLRRWRLEEKESVAGAQVTILLFGLGPG GGVTSAAASTLSEPPRTTQESSTRTRALGLPT LPMEKLAASTEPQGPRPVLGRSVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKKCRMCCCDVPAETLYDVLHDIE YKKKWDSNVEITFDIARLTVNADVGYYSWR CPKPLKRRDVITLRSWLEMGADVIIMYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLFKWVNKSSQFLAPKAMKK MYKACLKYPEWKGKHLPHFKFWLHIFEGSP LPSLALSELSVQHADSLENIDESAVAESREE RMGGAGGGGSDDDTSLYAEAPHRFFETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP							l
WVEEPVERNHOTELEVPRTRITTTOFAA QNLPNGYPRYPSGDASSHPSSRHPSVGSARL PSVGEESTHPILLVABEQVHTYYNTIGVQEER KNRTSYHVPLEARVSNAESSTPKEEPSSIEDR DPOILLEPEGVKEVLPPTVQKQLMEKEKLE QLGRQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASOVRGRITSTISTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMENYV NTENTYTASAIKKIEYSRRDCTFYTPHOEN RSSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSRKTRHNSTULPL 843 2193 A 6919 2 663 AGRPGTTHASGKNAYQSLREYLQIPPVSRA YTTACVLTTAAVQLELITPGLYFNDELIFKHF QIWRLITNELFGPGPFLFNMHELYRYCRM LEEGSFRGTADFVFMFLFGGFLMTLFGLFVS LVVFLGPGLYNNIGSSMCGABEPLCPHELLRP SQLPGPLSALGAHGFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMFIGGACGERME/FSLLFGLECNGVIL AHCNLRLPGSSNSPASASQVAGTGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPFWPDGGRFCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLRSQE GGVVTSAAASTLSEPFRRTQESRTRTALGLPT LPMEKLAASTEPQGFRPVLGRESVQVPDDQD FRSSRSECEAEVGMLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVEIFDIARLTVNADVGYYSWR CPKPLKNRDVTILRSWLPMGADVIMNYSVK HPKYPPRKDLVAAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYXACLKYPEWKQKHLLPHFKYPWLUPEQSP LPSLALSELSVQHADSLENIDESAVAESREE RMGGAGGGGOSDDDTSLYAEAPHRFETETIG PGAGRALGAAAAALSPLHPPGTWWHRARP		i					l
QNLPNGYPRYPSPGDASSHPSSRIPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRQVVGSGANTEWDTGVDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPFISLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTFTVPNFDIR RPSLEHRQLNYIQVDLEGGSDSDNQTPKTPT TPLPQTPTRTELYAVIDIERTAAMSNLQKAL PRDDGTSRKTRHNSTDLPL		ļ					
PSVGESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKPUGPTFVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASCVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLINGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTTPVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVUDIERTAAMSNLQKAL PRDDGTSRKKRHINSTUDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPQLYFNPBLIFKHH QIWRLITINFLFGPVGFNFLFNMHELYRVCRM LEEGSFRGRADFVFMFLFGGFLMTLFGLFVS LVFLGPGLYNN/GSSMCGAEEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWMKNSENTTYPESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERMEFSILLPGLECNGVIL AHCNLRJGSSNSPASASQVAGITGVCHHAR LIFVYSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLPFILLLPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDQD FRSFRSECAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YYRKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIMNYSVK HPKYPPRKDLVRASVLPMGADYIMNYSVK HPKYPPRKDLVRASVLPMGADYIMNYSVK HPKYPPRKDLVRASQVFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLHPEGSP LPSLALSELSVQHADSLENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWFHKARP	1		l	l I			
KNRTSVHYPILEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLINGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVPNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQPTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPPQLYFNPBLIFKHF QIWRLITINFLFFGPGFINTLFGLFVS LVFTLGPGLYNN/GSSMCGAEEPLCPHELLRP SQLPGFLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTTYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERMEFSLLLFGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTHPGYSFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGFRPVLGRESVQVPDDQD FRSTRSECCAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDLARLTVNADVGYYSWR CPKPLKNDVITTLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLHPEGSP LPSLALSELSVQHADSLENIDESAVAESREE RIMGGAGGGGSDDDTSLYABAPHFRETETG PGAGRALGAAAAAPALSPLHPPGTWWHRARP	1						1 -
DPQILLEPEGYKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLYYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLINGYHNNLDPMHNYV NTENVTYPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELY AVIDIERT AAMSNLQKAL PRDDGTSRIKTRHNSTULLP. 843 2193 A 6919 2 663 AGRFGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMLFGGFLMTLFGLFVS LVFLGPGLYNNIGSSMCGAEPPLCPHELIRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENITYTESYF 844 2194 A 6928 902 366 HRICMPOSSMCGAEPLCPHELIRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENITYTESYF HATCMPIGAGGERMEFSLLTGELENGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGKSQHTRPGYEFIPPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPPDDGGRPCVEMS TRAKKLRRIWRLEEKESVAGAVQTILLRSQE GGVTSAAASTLSEPPRRTQESRTRTALGLPT LPMEKLAASTEPQGFPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVTLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLIPHFKPWLHPEGSP LPSLALSELSVQHADSULPNIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHRRFETETG PGAGRALGAAAPALSSLHPPGTWVHRARP		į.					
QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLYENINGLSIPSASGVRRGRLTSTSTSD TQNRNNSAGRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTFTRRTELVAVIDIERTAAMSNLQKAL PRDDGTSR/KTRHNST\DLPL 843 2193 A 6919 2 663 AGRFGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFGPVGFNFLFNMIFLYRYCRM LEEGSTRGRTADFVFMFLEGGFLMTLFGLFVS LVFLGPGLYNN/GSSMCGAEEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHFFLGRCSQSTWWNKNSENTIYTESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTALGLPT LPMEKLAASTEPQGFRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDYPAETLYDVLHDIE YRKX WDSNYLFTDIARLTYNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPRYPPRRDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNRSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLWPEQSP LPSLALSELSVQHADSLENIDESAVAESREE RMGGAGGGGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	ŀ						ł .
SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPPLPQTPTPRTELYAVDIDERTAAMSNLQKAL PRDDGTSR\KTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYIQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L\VFLGPGLYNN\GSSMCGAE\EPLCPHELLRP SQLFGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMFIGGACGERMEFISLLFGLECNGYTL AHCNLRLPGSSNSPASASQVAGTTGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTTGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRTQESRTRTRALGLPT LPMEKLAASTEPQGFRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGTJLQSTGPKSCVIT YLAQVDPKGLAASTEPQGFRVLGSTGPKSCVIT YLAQVDPKGLASTEPWQADYLHFKPWLWPEQSP LPSLALSELSVQRADSLENIDESAVAESREE RIMGGAGGEGSLFWVVNKSSQFLAFKAMKK MYKACLKYPEWKQKHL\PHFKPWL\PPEQSP LPSLALSELSVQRADDTSLYALEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP						·	
TONINISAQRRTALLNYENLPSLPPWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENTYPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELTYAVIDLERTAAMSNLQKAL PRDDGTSRKTRHINSTUDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLITTAAVQLELITPFQLYFNPELIFKHF QIPKLITHFGPVGFNFLFMIFTLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS LVFLGPGLYNN/GSSMCGAERPLCPHELLRP SQLPGPLSALGAHGIPLVVGELNHCGPFGYCS WTHEFIGRCISQSTWWKNSENTITFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERMEFSLLIPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPFASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPRRTQESRTRTALGLFT LPPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKSLPKWVKNSSOFLAFKAMKK MYKACLKYPEWKQKHLVPHFKPWLWPPEQSP LPSLALSSELSVQHADSLENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	ļ	ļ					
LŚRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTFTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPOTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTIACVLITAAVQLELITIPPQLYFNPELIFKHF QIWRLITINFLFFGPVGFINIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS LVFLFGGLYNNIGSSMCGAEEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERMEFSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLPPPWPDGGRPCCVEMS TRAKKLRIWRILEKEKSVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADVIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLHPEQSP LPSLALSELSVQHADSSLENIDESAVAESREE RWGGGAGGGGSDDDTSLYAEAPHBFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP		•		1			
843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPYSRA YTTACVLTTAAVQLELITPFQLYFNFELIFKH QIWRLITINFLFFGPVGFNFLFNMFLJYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS LVFLGPGLYNNIGSSMCGAEEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGTIGSCHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGVYTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRILHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDLARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLLQSTGPKSCVIT YLAQVDPKGSLPKWVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLHPEQSP LPSLALSUELSVQHADSVLENDIDESAVVAESREE RIMGGAGGGGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP							TQNINNSAQRRTALLNYENLPSLPPVWEARK
RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSR\KTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYENPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS LVFLGPGLYNN/GSSMCGAE\(\text{PLCPHELLRP}\) SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLPFILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLFT LPMEKLAASTEPQGPRPVLGRESVQVPDQD FRSFRSECEABVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLYRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\(\text{PLSVQHADS\(\text{LENDESA}\) RMGGAGGGG\(\text{GSDDDTSLYAEAPPHFRFETETG} PGAGRALGAAAAPALSPLHPPGTWWHRARP	ŀ		!			•	LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV
843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGFLMTLFGLFVS LAVFLGPGLYNN/GSSMCGAEEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTTYFESYF HACALRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRTQESRTRTRALGLPT LPMEKLAASTEPQGFPPVLGRESVQVPDDQD FRSFRSECEAFVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVTLRSWLPMGADGTINNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKOKHLPHFKPWLWPEQSP LPSLALSELSVQHADSLENIDESAVAESREE RUGGGGGGGGGDDDTSLYAEAPHFRFETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	ł	i	ŀ				NTENVTVPASAHKIEYSRRRDCTPTVFNFDIR
843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFGPVGFNILFNMIFLYRYCRM LEEGSFRGRTADFVFNMIFLFGFLMTLFGLFVS LVFLGPGLYNN/GSSMCGAEEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTTYFESYF HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTILLLRSQE GGVVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMPSVK HPYKYPPKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLFKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLWIPEQSP LPSLALSLSVQHADSLENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP		ŀ					RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT
843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFGPVGFNILFNMIFLYRYCRM LEEGSFRGRTADFVFNMIFLFGFLMTLFGLFVS LVFLGPGLYNN/GSSMCGAEEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTTYFESYF HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTILLLRSQE GGVVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMPSVK HPYKYPPKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLFKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLWIPEQSP LPSLALSLSVQHADSLENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP		ĺ		[TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL
843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLITAAVQLELITPFQLYFNPBLIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE:PLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTTYFESYF HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVVTSAAASTLSEPPRRTQESRTRTRALGLFT LPMEKLAASTEPGPRRVDGSRTSTRALGLFT LPMEKLAASTEPGPRRVDGSRTSTRALGLFT LPMEKLAASTEPGPRVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKK WDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLLQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLWIPEQSP LPSLALSVQHADS\LENIDESAV\AESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG G PGAGRALGAAAAPALSPLHPPGTWWHRARP	1				,		
YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\PLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPJQGACGERME/FSLLLPGLECNGVIL AHCNLRJGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKRIWRILEEKESVAGAVQTLLLRSQE GGVYTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLHPBQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	843	2193	A	6919	2	663	
QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAEVEPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEKESVAGAVQTILLRSQE GGVYTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLWIPPEQSP LPSLALSVQHADSVLENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP		1			_		
LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAEVELCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGV\TSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKK WDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYXACLKYPEWKQKHLVPHFKPULYPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP		1					, , ,
L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERMEFSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTLEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLLQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLPHFKPWLHPEQSP LPSLALSVELSVQHADSULENIDESAVVAESREE RMGGAGGEGGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	ľ	i i				İ	
SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF 844 2194 A 6928 902 366 HRLCMPIQGACGERMEFSLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLJQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSVELSVQHADSVLENTDESAVAESREE RWGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP			Ì	i I			
844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKK WDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLWPEQSP LPSLALSÆLSVQHADSVLENIDESAVAESREE RVMGGAGGEGVSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP				!		·	
844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVITSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKK WDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLWPEQSP LPSLALSVELSVQHADSVLENIDESAVAESREE RWGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP							
AHCNLRLÞGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSELSVQHADSVLENIDESAVAESREE RVMGGAGGEGVSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	944	2104	├ ~	6020	002	366	
LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVYTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSELSVQHADSVLENIDESAVAESREE RVMGGAGGEGVSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	044	2174	^ ^	0928	902		
QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM 845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVTSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSELSVQHADSVLENIDESAVAESREE RIMGGAGGEGSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP			J	J			
845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGVITSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP			.				•
845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGV\TSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP			1			·	
TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGV\TSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP			L				
GGV\TSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	845	2195	Α	6939	1660	317	
LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSIELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP			1				
FRSFRSECAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSYELSVQHADSVLENIDESAVAESREE RVMGGAGGEGVSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP		([
FRSFRSECAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSYELSVQHADSVLENIDESAVAESREE RVMGGAGGEGVSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	1		·				LPMEKLAASTEPQGPRPVLGRESVQVPDDOD
EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSYELSVQHADSVLENIDESAVVAESREE RVMGGAGGEGVSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP							FRSFRSECEAEVGWNLTYSRAGVSVWVOAV
YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALSYELSVQHADS\LENIDESAV\AESREE R\mgGAGGEG\SDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP			1				
CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP	ļ			1			
HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP							
YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP							
MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP		•					
LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP							
R\MGGAGGEG\SDDTSLYAEAPHRFRETETG PGAGRALGAAAPALSPLHPPGTWWHRARP					İ		
PGAGRALGAAAAPALSPLHPPGTWWHRARP]			ľ			
{						'l	
	L			<u> </u>			KKVLQPGWTEPQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
846	2196	A	6944	sequence 42	2672	nucleotide insertion RRKMAGCRGSLCCCRWCCCGERETRTPE ELTILGETQEEEDEILPRKDYESLDYDRCINDP
						YLEVLETMDNKKGRRYEAVKWMVVFAIGV CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE AGSGITEGKCYLYARQVPGLVRLPTLLWKAL GVLLTVAAMLLI\GLGSPMIHSGSVVGAGLPQ FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV VMGVIGGLLGATFNCLNKRLAKYRMRNVHP KPKLVRVLESLLVSLVTTVVVFVASMVLGEC RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESAILQLFHQDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA AFLGGVVRMTISLTVILIESTNEITYGLPIMVT LMVGKWTGDFFNKGNYDIHVGLRGVPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQQMLERRYTPY PNLYPDQSPSEDWTMEERFRPLTFHGLIRSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE
847	2197	A .	6951		1994	IVGIITRHNLTYEFLQARLRQHYQTI NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVFQEK WERA YFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIK WLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM
849	2199	A	6999	963	5	LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV VSYMCILCAILQIQSREVQRKAFCTCFSHLCVI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isolaucine, K=Lysine, L=Laucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide	i	/-possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
						GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS
L	0000		5001			LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL
850	2200	A	7001	1	.1011	MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI
Į.						DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG
}						KVARRRVGATWLLHLAVADLLCCLSLPILAV PIARGGHWPYGAVGCRALPSIILLTMYASVLL
<u> </u>		ļ				LAALSADLCFLALGPAW\CLRFS/GACGVQVA
1		1	i			CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ
1		ĺ	i 1			CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA
						SCHSALLCWAARRCRPLGTAIVVGFFVCWAP
					•	YHLLGLVLTVAAPNSALLARALRAEPLIVGL
1		İ				ALAHSCLNPMLFLYFGRAQLRRSLPAACHW
1			1			ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	A	7011	1	2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI
						SRGVLVCDECCSVHRSLGRHISIVKHLRHSA
			i l	{		WPPTLLQMVHTLASNGANSIWEHSLLDPAQV
						QSGPALKQTPKDKV\HPIKSEFIRAKYQMLAF
						VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE
						TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG
				J		QTLQAELLVVYGADPGSPDVNGRTPIDYARQ
			1			AGHHELAERLVECQYELTDRLAFYLCGRKPD
1						HKNGHYIIPQMADSLDLSELAKAAKKKLQAL
						SNRLFEELAMDVYDEVDRRENDAVWLATQN HSTLVTERSAVPFLPVNPEYSATRNQGRQKL
						ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD
						NLELSLRSQSDLDDQHDYDSVASDEDTDQEP
						LRSTGATRSNRARSMDSSDLSDGAVTLQEYL
				,		ELKKALATSEAKVQQLMKVNSSLSDELRRLQ
						REIHKLQAENLQLRQPPGPVPTPPLPSERAEH
		}				TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG
1.						PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL
i						YRIRKGVSASAVPFTPSSPLLSCSQEGSRHTSK
1						LSRHGSGADSDYENTQSGDPLLGLEGKRFLE
1						LGKEEDFHPELESLDGDLDPGLPSTEDVILKT
	'					EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA
						VTEMASLFPKRPALEPVRSSLRLLNASAYRLQ
						SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA
852	2202	A	7016	484	1777	RAAKQLVTITTREKKQ
0.72	2202	A	1,010	704	1///	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL TVKGLLKPSFSPRNYKALSEVQGWKQRMAA
						KELARONMOLGFKLLKKLAFYNPGRNIFLSP
	;		,			LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP
			(EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID
						QRLQPQRKFLEDAKNFYSAETILTNFONLEM
						AQKQINDFI/ESKTHGKINNLIENIDPGTVMLL
.						ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS
1						VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK
						NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL
						LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS
1						KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM
						DERGTEGAAGTGAQTLPMETPLVVKIDKPYL
963	2000		70:-		2002	LLIYSEKIPSVLFLGKIVNPIGK
853	2203	A	7017	1 .	3293	MTHACNPSTLGGQGRRITRSHGRRRSSRGPV
		-				ARHVAAGAGHENKHGGSRRFPAGVAPRRAM
1]				ANVSKKVSWSGRDRDDEEAAPLLRRTARPG
					ł	GGTPLLNGAGPGAARQSPRSALFRVGHMSSV ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLKY
				·		ESLDYDNSENQLFLEEERRINHTAFRTVEIKR
						WVICALIGILTGLVACFIDIVVENLAGLKYRVI
	i,			l	ł	KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH
						VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG
						NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI AMGVVGGVLGAVFNALNYWLTMFRIRYIHR PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL QGGSMŞYPLQLFCADGEYNSMAAAFFNTPEK SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT
		:				YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG AAIWADPGKYALMGAAAQLGGIVRMTLSLT VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV MSTPVTCLRRREKVGVIVDVLSDTASNHNGF
					·	PVVEHADDTQPARLQGLILRSQLIVLLKHKVF VERSNLGLVQRRLRLKDFRDAYPRFPPIQSIH VSQDERECTMDLSEFMNPSPYTVPQEASLPR VFKLFRALGLRHLVVVDNRNQVVGLVTRKD LARYRLGKRGLEELSLAQTGPKAQATAEGRV
			•			AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP LSLEELSERYESSHPTSTASVPEQDTAKHWNQ LEQWVVELQAEVACLREHKQRCERATRSLL RELLQVRARVQLQGSELRQLQQEARPAAQAP EKEAPEFSGLQNQMQALDKRLVEVREALTRL
						RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ QEEQGREVACGALQKNQEDSSRRVDLEVAR M
854	2204	A	7037	139	2604	AGTWEPRPYDQAKETGAPGSQPPVPPMELRP WLLWVVAATGTLVLLAADAQGQKVFTNTW AVRIPGGPAVANSVARKHGFLNLGQIFGDYY
				ī - -		HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL EQQVAKRRTKRDVYQEPTDPKFPQQWYL\SG VTQ\RDLMVKAAWAQGYTGHGIVVSILDDGI EKNHPDLAGNYDPGASFDVNDODPDPOPRY
						TQMNDNRHGTRCAGEVAAVANNGVCGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNPN HIHIYSASWGPEDDGKTVDGPARLAEEAFFR
						GVSQGRGGLGSIFVWASGNGGREHDSCNCD GYTNSIYTLSISSATQFGNVPWYSEACSSTLA TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS APLAAGIIALTLEANKNLTWRDMQHLVVQTS
						KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY
						SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPOVLDTHYSTENDVETIRASVCAPCHAS
						CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWOEECPSD
055	2205		7050		1441	SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	1441	QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL EDGNFLKLANNGTVLRASHGTKMMTPEVLA
Ll	l					EAYGKKEWKHFLSDTGMACRSGKYYFYDN

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVYYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFF\IDSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVTYPS MCWIRDSLVSYTTNIGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLG\LP WALIFFSFASGTFQLVVLYLFSITTSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	A	7136	3 .	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
					·	HPPPEEDQGEERPRLWVMPNHQVLLGPEEDQ DHIYHPQ*GSRGHHCPRPVPRPRLLGLGPSLP CPS
861	2211	A	7161	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI LTTTTRIMFHCTYLFASVCLSLLNTLLSPNCL KSAMILQ
862	2212	A	7211	665	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF YKNIQKLSFSNYVYHQNYVFSSDWSYDF
863	2213	A	7212	924	1273	HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ CIKPNYQSPPKECDYNILANSVA
864	2214	A	7214	845	1619	SDKGGKKADRKNHLRHAFPLLPHRVRERLH DPKVPVDADHVQGQDPGRAAHDIHGEDVTE KVSKDPLAPDEVGDTDEGHDRHGHREVGQR HGHDQEEVAYEERACEGGKFATVEVTDKPV DEALREAMPKVAKYAGGTNDKGIGMGMTV PISFAVFPNEDGSLQKKLKVWFRIPNQFQSDP PAPSDKSVKIEEREGITVYSMQFGGYAKEAD YVAQATRLRAALEGTATYRGDIYFCTGYDPP MKPYGRRNEIWLLKT
865	2215	A	7246	559	682	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS LANMAKPRLY
866		Α	7257	641	1310	TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL DLKKSDFSTRWQKQRCPVVKSKCRENASPFF FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV QIPLTESYCGPCPKNWICYKNNCYQFFDESKN WYESQASCMSQNASLLKVYSKEDQDLLKLV KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT IIEMQKGDCALYASSFKGYIENCSTPNTYICM QRTV
867	2217	A	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI MPFFQTLWLMNANRFCSIFTTTNVANNCWW TPYHCWLSVVVCRCESHGI
868	2218	A	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP KGSCPAGGSRMVSESD*EGRGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP
869	2219	A	7332	1223	332 .	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWILLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHILLA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLETGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF
870	2220	A	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA
871	2221	Α	7403	3	393	SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR ALRGAALPGESEAGDPESLRSSVNADWIQYS

Deptide Control Cont	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Uence							
1914 ng to first amino aed residue of peptide pequence peptide pequence peptide s	cotide	seq-		USSN	location		
mino acid residue of peptide residue of peptide residue of peptide sequence T-Threonine, V=Vallne, W=Typtophan, Y=Tyrosin, X=Unknow, X=Stop codon, P=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide insertion DLWRAEVSTPRCEAGFCQECFRTPGNQEKDQ		uence					
	uence			914			
Peptide Sequence	į.						
		1				scquate	
DI.WEAEVSTIRCEAGFCQECERTFGNQEKDG PRIC		ļ	}				
\$72 2222 A 7413 1061 359							DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG
PGGS*PQATI.HI.DRARVSAŠPTK.EIQVAKVVOFIMCRED RMIMSPYKNNVGROLNIALVNOTITGAVLOG KARDMYSGDVMIL VKFILKEPGGALVLVAS YDDPGTKAMDESRKLFSDLGSSYAKOLGFGALVLVAS YDDPGTKAMDESRKLFSDLGSSYAKOLGFGALVLVAS YDDPGTKAMDESRKLFSDLGSSYAKOLGFGALVLVAS SVPUFGAKDLRGKSPTEOFLKEQPQTONKYE GWPEILEMEGCMPPKDF RTS				l			
CGLIKPĆPANYTAFKICSGALANVVOPTMCEN RMMSPVKNNYGGIMALVSGITGAVLOQ KAPDMYSGDVMHLVKFILKEPGGALVLVAS YDDPOTKMMDESKLISDIGSSVAKOLGFRD SWYFIGAKDLRGKSPFEOFLKEQPOTONKYE GWPELLEMEGCMPFKPF RKCAGHIGSSCL SQHFGRLRWEDRLRLGVQ DHPGGHCSTPSLLKIERKLF A 7468 146	872	2222	A	7413	1061	359	
RMMSPYKNYGRQLNIALVOTTGAVLOG KAPDMYSGDVMILVFIKERPGGALVLVAS YDDPGTKMDESRKLFSDLGSSVAKQLGFGALVLVAS YDDPGTKMDESRKLFSDLGSSVAKQLGFGALVLVAS YDDPGTKMDESRKLFSDLGSSVAKQLGFGALVLVAS YDDPGTKMDESRKLFSDLGSSVAKQLGFGALVLVAS SWYFIGAKDLRGKSPFEOFLKEQPQTONKYE GWPELLEMEGCMPFKFF 873 2223 A 7429 2242 2394 ILKCAGHIGGSCLFSQHFGRLRWEDKLRLGVQ DHPGQHCETTSLLKERKLF 874 2224 A 7468 146 894 PCTSCVLWATLHLPASTRKAPQAECGMISITE WQKIGVGTTGGTGFLLGTLLYTDSVLLAFGN LLFLTGLSLIGLRKTFWFFORHKLKGTSLLAFGN LLFLTGLSLIGLRKTFWFFORHKLKGTSLLAFGN LLFLTGLSLIGLRKTFWFFORHKLKGTSLFFPV AFGFLGNVCNIPFLGALFRRLQGTSSMV*KLGGVVLLRWPLLGMFLTCHFFORHKLKGTSLFFPV AFGFLGNVCNIPFLGALFRRLQGTSSMV*KLGGVVLLRWPLLGMFLTCHFFFLFKGFFPV AFGFLGNVCNIPFLGALFRRLQGTSSMV*KLTSMSLDHWLKGAKREEWEPPGSPALTHSE TYPGPPQVQKERNGASQLTSNPQVDSRGGQE AEMQTPRILQGWVHILTLV WEEK 875 2225 A 7498 91 251 GEKPVPTWLQDEAGQWLLGFVAQPWGWF SRRHEP*HGGVLFLLGFSAAPPGKL 876 2226 A 7544 403 587 YSCLCFLFKHTISTNSVHIWLGTVVHAYNPN LLGQGGGWLAGQFEKTSLGVVHAVAPPN LLGQGGGWLAGQFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGWA-QGFEKTSLGVVHAVAPPN LLGQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG				<u> </u>			
		Į.		}		i	
STO 2223 A 7429 2242 2394	1	i					•
ST 2223 A 7429 2242 2394 II.K.CAGHGGSCL *SQHFGELRWEDRLRIGVQ DHPQHCETPSLLKIERKLF FOR SCHOOL ** Comment of the comment of the	1			<u> </u>		ŀ	
873 2223 A 7429 2242 2394]	}]			
DHPGGHCETPSILKÜERKLF							
874 2224 A 7468 146 894 PCTSCVLWATLHLPASTRKAPQAECGMISTIF	873	2223	Α	7429	2242	2394	
WQKIGVGITGFGIFFILFGTILLYEDSVILAFGN LLFLTGLSLIGRKTFWEDRIKLKGTSPIL GGVVTVLRWPLLGMFLETYGFTSLFKGFFPV AFGFLGNVCNIPFLGALFRRLQGTSSMV-KITE MSSLNLDHWLKGAKREEWEPPPQSPALTHSP TYFCPPQVQKERNGABQLTSNPQVDSRGCQE AEMQTPRRLGWGWYHTLTVLWEEK AEMQTPRLGWGWYHTLTVLWEEK AEMQTPRLGWGWYHTLTVLWEEK AEMQTPRLGWGWYHTLTVLWEEK AEMQTPRLGWGWYHTLTVLWEEK AEMQTPRLGWGWYHTLTVLWEEK AEMQTPRLGWGWYHTLTVLWEEK AFGFLGGGGWA*GQFFKTSLGNTVRQCYNGWGWPG SERHEP*HGGVLFRLGPSAPPGKL SFR	874	2224	_	7468	146	804	
LLĒ.TGLS.LIGI.RKITFWFFFQRKILKGTSFLL GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE MSSLNLDHWLKGAKREEWEPPPQSPALTHSP TYPGPPQVQKERNGAGQLTSNPQVDSSGCQE AFGFLGNVCNIPFLGAGEQLTSNPQVDSSGCQE AFGFLGNVCNIPFLGAGEQLTSNPQVDSSGCQE AFGFLGNVCNIPFLGAGEQLTSNPQVDSSGCQE AEMQTPRRLGWGWYHTLTLYLWEEK B75 2225 A 7498 91 251 GEKPVPTVLQDEAGQLTSNPQVDSWGPG SERHEP*HGGVLFRLGPSAPPGKL AFGFLGNVCNIPFLYBEGGGAFWVALVARGGC AFGFLGNVCNIPFLYBEGGGAFWVALVARGGC AFGFLGNVCNIPFLYBEGGGAFWVALVARGGC AFGFLGNVCNIPFLYBEGGGAFWVALVARGGC TFKDKVLVAARRNASVLVINEERYCNITLP MSHAGTGNIVVINISYPKGREILELVQKOIPV TMTIGVGTRHVQETISGQSVVEVALFITTMIII SLAWLIFYYIQRFLYTGSQIGSOSHRKETIKKVI GQLLLHTVKHGEKGIDVDAENCAVCENFKV KDIRRLPCKHIFHRICITYSQIGSOSHRKETIKKVI GQLLLHTVKHGEKGIDVDAENCAVCENFKV KDIRRLPCKHIFHRICITYSQIGSOSHRKETIKKVI GQLLLHTVKHGEKGIDVDAENCAVCENFKV KDIRRLPCKHIFHRICITYGSQIGSOSHRKETIKKVI GQLLLHTVKHGEKGIDVDAENCAVCENFKV KDIRRLPCKHIFHRICITYGSQIGSOSHRKETIKKVI GQLLLHTVKHGERGIDVDAENCAVCENFKV KDIRRLPCKHIFHRICITYGSQIGSOSHRKETIKKVI GQLLLHTVKHGERGIDVDAENCAVCENFKV KDIRRLPCKHIFHRICITYGSQIGSOSHRKETIKKVI GQLLLHTVKHGERGIDVDAENCAVCENFKV KDIRRLPCKHIFHRICITYGSQIGVSUSV KDIRRLPCKHIFHRICITYGNITYCTANILTON GKSELLKSGSKSTLKHIWTESSKDLSISRLLS QTFRGKRONTONLDVERPPYEPSSG STEWILKWWSPTCLDQLNGNGTFSVYT RHNSTGQGNVSVSL VPPFKVEFDLAQQTVID AKDSKSPNCREEYEKVDKATKNTLCNVDPSK TESWKLKWWSPTCLDQLNGSAPGNVFHG GGSRGSBRGSGGGOSLYPVYLDRQVPDTS VGLTDRLIVKRCWDIALGFLAGPMNFHIY MAGNTISIPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLDRQVPDTS VQETDRLIVKRCWDIALGFLAGPMNFHIY MAGNTISIPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAHGERERSGGGL LL SPPCTMRSHITIMTTTSVSSWPYSSHRMRFIT NHSBQPPONFSATPNVTTCPMDEELLSTVLTT SYSVETVGLVONIALJAGGFLYMINGIN CKVVGTLPYMNMTSILLGFLIBLRYIKINGST CCCCCC AFGRAND CCCCCCC AFGRAND SERMINT SYSVETVA CCCCCC AFGRAND CCCCCC AFGRAND CCCCCC AFG	874	2224	^	7400	140	654	
### AFGFLGNYCNPFLGALFRRLQGTSSMV*KTE MSSLMLDHWL KGAKEWPPPQSSAL 1HSP TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE AEMQTPRALGWGWYHTLTLYLWEEK GEKPVPTVLQDEAGQWLLGFVAQPWGWPG SERHEP*HGGVLFRLGPSAPPGKL 7500 S	}	Į		}			
MSSLNLDHWLKGAKREEWEPPPQSPALTHSP				į		1	GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV
TYPGPPQVQKERNGABQLTSNPQVDGRGCQE			1	[
AEMQTPRR_GWGWYHTLTLYL_WEEK							·
875 2225 A 7498 91 251 GEKPVPTWLQDEAGQWLLGFVAQPWGWFG 876 2226 A 7544 403 587 YSCLCFLFKHITSKNSVHWLGTVVHAYNFN 877 2227 A 7566 2 940 GCAPDTRFFVEPGGGAAPWVALVARGC TTKDKVLVAARNASAVLYNEERYGNITLP MSHAGTGNIVVIGUSGGSAFKETKKVI GQLLHTVKHGEKGIDVDAENCAVCLENFKV KDIRILPCKHIFRICIDPWLLDHTCPMCKL DVIKALGYWGEFGDVGEMPAPESPPGRDPAA NLSLALPDDDGSDESSPFSASPASEDEQCDPSF KGDAGENTALLEAGRSDSRHGGFIS 878 2228 A 7586 315 1232 ERSLLCKVDVRWITVSEGTKTQRTHRQGSLR RGRMQACWYLFLLQFIVVLYCANLING GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTFEYSEQDLWDW LRNSTDLQEPFRAKRAPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKVDHGNGTFSVF RHNSTGGGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFFSG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANGGRFKWAIELSGPG GGSRGSDGGSGDSLYPYGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMILFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAYYK CQSMGLLPTHASDWLAFIEPFERMEFSGGGL LL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWFYSSHRMRFIT NISDQPPQNFSATPNVTTCMDEKLLSTVLTT SYSVIFIGLYGNIALLYFIGHRKNNSIQIYL LNVAIADLLLIFCLPFRMYHINQNKWTLGVIL CKVVGTLFYMNYISILLFGISLDRYIKINGSI QQRKAITTKQSIVVCCIVWMLALGGFLTMILL							
876 2226 A 7544 403 587	875	2225	A	7498	91	251	
ILGGQGGWIA*GQEFKTSLGNTVRPCLYK							
877 2227 A 7566 2 940 GCAPDTRFFVPEPGGRGAPWALVARGGC	876	2226	Α	7544	403	587	
TFKDKVLVAARRNASAVVLYNEERYGNITLP MSHAGTGNIVVIMISYPKGREILELVQKGPV TMTGVGTRITVQEFISGQSVVPVAIAFTMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGIDVDAERCAVCIENPKV KDIRILPCKHIEHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS 878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQROVSVSLVPPTKIVEFDLAQQTVIU AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTYFFSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPFERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALLYVFLGHRKRNSIQIYL LNVAIADLLLIFCLFFRMYHNVQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	000	2002		5566		0.10	
MSHAGTGNIVVIMISYPKGREILELVQKGIPV TMTIGVGTRHVQEFISQSVVPVAIAFITMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIRRLPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYGEFGDVQEBPGBVQESPPSGRPAG NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS 878 2228 A 7586 315 1232 ERSILCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSKSTLKHHWTESSKDLSISRILS QTFRGKENDTDLDLRYDIPPEPYSEQDLWDW LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTQQNSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWSPTCLDQLNGSAPGNVFTHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPC GGSRGRSDRGSGQDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDDIALGPLKQUPMILFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLINGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMSHTITITTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALLYVFLGHRKRNSIQIYL LNVAIADLLLIFCLFFRMYHNQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITIKQSIYVCCIVWMLALGGFLTMIIL	8//	2221	A	/300	2	940	1
TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII SLAWLIFYYIQRTLYTGSQIGSQSHREETIKKVI GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAFESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGFIS 878 2228 A 7586 315 1232 ERSILCKVDVRWIYVSEGTKTQRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRILS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITIGKVDHGNGTFSVF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNVDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGLVANRGRFKWALEJSGPG GGSRGRSDRGSQGODSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFTEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKRNSI QQRKAHTKQSIYVCCIVWMLALGGFLTMIIL			į				
GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS 878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDQDENPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTTSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTIMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL LL LL LL LL LL LL							
KDIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRPAA NISALALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS RGPAGENTALLEAGRSDSRHGGPIS RGPAGENTALLEAGRSDSRHGGPIS RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLR YDTPEPYSEQDL WDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTDQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNVDPSK TCYQEQTQSHVSWLCSKPFK VICIYISFYSTD YKLVQKVCPDYNYHSDTTYPFPS TCYQEQTQSHVSWLCSKPFK VICIYISFYSTD YKLVQKVCPDYNYHSDTTYPFPS GGSRGSGGGVSLYVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQPMNLFIMY MAGNTISIPTIMVCMMAWRPIQALMAISAT FKMLESSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL SPQKTMRSHITIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADILLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMMYNSIILLGFISLDRYIKNRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL				İ			
DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS 878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRFKWAIELSGPG GGSRGRSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHBNOKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL			ľ				•
NLSLALPDDDGSDESSPSASPAESEPQCDPSF KGDAGENTALLEAGRSDRHGGPIS 878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLR YDTFEPYSEQDL WDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIETVGLVGNIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHNQNKWTLGVL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
KGDAGENTALLEAGRSDSRHGGPIS							
RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLR YDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTIMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHITIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFVGLVGNIALLYFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL			ļ	,			
GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLR YDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQYL LNVAIADLLLIFCLPFRIMYHNQNKWTLGVIL LNVAIADLLLIFCLPFRIMYHNQNKWTLGVIL CKVVGTLFYMNWYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	878	2228	Α	7586	315	1232	
RESIDENT A TOTAL OF THE PROPERTY OF THE PROPER							
LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTL.CNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	1						•
RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL		· ·	İ				
TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTILFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	1						RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	1						
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLYYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL						<i>'</i>	
880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	870	2220		7605	470	301	
GGSRGRSDRĞSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVTL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
VQETDRĪLVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL				'			
FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL		,					VQETDRILVEKRCWDIALGPLKQIPMNLFIMY
CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	881	2231	A	7615	291	1452	
LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL				_			NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT
CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	ŀ						
							TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN SGKYATTARNSFIVLIIFTICFVPYHAFRFIYISS QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP
882	2232	A	7617	67	379	VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST SEFKPGYSLHDTSVAVKIQSSSKST RQMALLKANKDLISAGLKEFSVLLNQQVFND PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	A	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK
884	2234	A	7638	2640	2861	AVATVGPISVAVGASHVFFQFYKKGKHLSS APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP GLSCHTSHSG
885	2235	A	7642	201	455	PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	A	7702	242	1298	APSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707	185	2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	-	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ł			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	_	/=possible nucleotide deletion, \=possible
L				sequence	l	nucleotide insertion
	1				T	EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR
	ŀ					GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP
					1	DESTKTKDQILTSRINAVERDLLEPSPADQLG
ľ			1		1	NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV
		ŀ	l 1			TREPARRLFLFGEEPSKLDQDVLAALECADV
	1	ļ				DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV
1 .					i	KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP
	22.12	<u> </u>				GLGSPGRYSPVHGSQLRRMARLAELAAL
890	2240	Α	7711	360	269	RHMPVIPALWEAEVGGLLEPRSSRSAWATE
891	2241	Α	7721	61	1175	KLPWEPSFLIKMQURHSEQTLKTALISKNPVL
]						VSQYEKLDAGEQRLMNEAFQPASDLFGPITL
						HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN
]						KRSIYIQSIGSLGNTRIISEEYIKWLTGYCKAYF
1				,		YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH
j .						AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS
						WNFVFGQASLTDGVGIFSFARYGSDFYSMHY
1		ļ				KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL
						EEADRRPLNLCPICLHKLQCAVGFSIVERYKA
		}				LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
!						EAFKEWKEWIIKCLAVLOK
892	2242	Α	7723	2	1650	SAPTAPARPCRAERGSGGGMLALLAASVALA
"						VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL
l i						PAMPMQGGAQSPEEELRAAVLQLRETVVQQ
1						KETLASARAIRELTGKLARCEGLAGGKARGA
						GATGKDTMGDLPRDPGHVVEQLSRSLQTLK
1				·		DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ
1						LRETVVQQKETLASARAIRELTGKLARCEGL
l i					<i>(</i> ·	AGGKARGAGATGKDTMGDLPRDPGHVVEQ
1 1	1			'		LSRSLQTLKDRLESLEHQLRANVSNAGLPGD
						FREVLQQRLGELERQLLRKGAELEDEKSLLH
1						NETSAHRQKTESTLNALLQRVTELERGNSAF
1 1						KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT
	l					ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE
				.		WGNNPIELLINDKVAQLPLFVSDGKWHHICV
[[• 1	TWTTRDGMWEAFQDGKKLGTGENLAPWHPI
) 1						KPGGVLILGQEQDTVGGRFDATQAFVGELSQ
						FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
902	2242	<u> </u>	7720	2554	2410	NNVDVFGGASKWPVETCEERLLDL
893	2243	A	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL
						DNYNDTSLVENHLCPATEGPLMASFKAVFVP
						VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET
1	ļ]				FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF
				l	1	LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV
] .				l		HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI
1 1	ļ	l		ļ	,	LFAKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR
] [İ		QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV
			1	!	ł	IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL
	ļ	}	1	1	j	GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
			ŀ			CTGPASLCQLFPSWRRSSLSESENATSLTTF
894	2244	A	7738	670	287	FVTRAGRWGAGARVRGGAGGMASGAARWL
"			1130	3,0	201	VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR
	. [.]			SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI
	Į	ļ				VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
	j	l	ł	ì		D D
895	2245	A	7753	119	278	APYAHSQVHCLDKVCGLLPFLNPEVPDOFYR
		·-				LWLSLFLHAGKEAPHCPRTRPL
896	2246	A	7754	1	372	SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM
		لـــــــــــــــــــــــــــــــــــــ			-:	

nucleotide seq uence USSN 09496 puence politic form of the property of the pro	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence USSN professories of state and professories of state and professories of peptide peptide sequence pept	NO: of		hod				D=Aspartic Acid, E=Glutamic Acid,
Sequence 914			1			I .	
Bene				1			
### ### ### ### ### ### ### ### ### ##	, .	ucnce	1	1			
Pepside Peps	nence		Ì	914			
Poptide Popt							
	ł	ł	ļ			sequence	
897 2247 A 7761 1725 445 PRERGRYCADPSEWOKYVSDLELSA	1	1	ł	ĺ			
## ACCFSYTSRQPONFIADVETSSQCSKPSVIFL TKRGRQVCADPSEEWVGKYVSLELSA ## TKRGRQVCADPSEEWVGKYVSLELSA ## RPRREGTHIFSCVLGSFRVSAMPFRVSTFLFL RPSREPILSGSPETSAAAMILITVRHGTVRY RSSALLARTKNINQRYFGTNSVICSKKDXGSV RTETSKETSESQDSEEKDTKXDLLGIKGMK VELSTVAVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIBELSQL QQHEESGRAQRDAKRPKSSNI ISDMKVARSATARVRSRPELRQDEGYDNYP GQRETIDDLKKKRNIFTGKRINFDMMAVTKE APETDTSPSL WUPEAKQLATVNEQPLQNGF BELQWYREGGK LEPRINRAGFDDDGSEPH EHILEKHLESPRQOPRHEMELVTCCLSKNP TLYSVKQKVEHEWVRTYNEKDLLKESNIQ KLRPWKRLFRNN ## SOOTTOPPAGSCSTGMRMILLFTAILAFSLA QSFGAVCKEPQEVVPGGGRSKRDPDLVQLL QRLFKSHSLEGLLKAJQASTDPKESTSPEK RDMEDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKVPPAS ## SOOTTOPPAGSCSTGMRMILLFTAILAFSLA QSFGAVCKEPQEVVPGGGRSKDPDLVQLL QRLFKSHSLEGLLKAJQASTDPKESTSPEK RDMEDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKVPPAS ## STANKSSLGGLKALSQASTDPKESTSPEK RDMEDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKVPPAS ## SFSMRSSKNAVGKEPMMARAQLGRCQLIM QSEMRERQMAMQLAWSEELKYFGTFFGLA ASLTAGAKKKKPAFLVPPUS-SELT-TVOYDL GVGTILLERMKGEADDLISTEKSKLQLPRGMT FSSIKARKSQSRFENV ## STANKSKAJQAKTPKYTY QFTASKARGSGRFENV ## STANKSKGPGSAVSPYFTFMSSDVAALH KAMVKGVDEATIDDLITKRNNAQRQQIKAAY LQTGGKRASTQTHETSTAYALKT KERDINFNYREELKRICHHEEVYLALLKTP AQFDADELRAAMKGVGTHALALIKTP AQFDADELRAAMKGVGTHALALIKTP AQFDADELRAAMKGVGTHALALIKTP AQFDADELRAAMKGVGTHALALIKTP ALSLAKGDRSEDFFVNDLADSDARALYEA GERKGTONVYPTILTTRSTYPQLRRYFQKY TKYSKEDMNDIKAFYQKMYGISLQAALDETKGD PKILVALCGGN ## STANKSKGRGGAVSEKYPULJQQLOVKIRHCKGER ## QAPPASELTRIGIGRRCTATINATOCAUPHIL ADLEQRLPKADQLERSGLNEDLEKLQMARP NILGLERNAYCMAQLLDNSDTAEPTKAGROA SQPTTTPASDAPGRKLEGCRFHGYHRRMH SVGRYFSKWGSSPNRSRRHSPHQALRKGVRR ARROKGTAARRQKGTA					Sequence	 	
				i .		İ	ACCESTED OFFICE A DALLES OF THE ACCES A DESCRIPTION OF THE ACCES AND ACCES A DALLES OF THE ACCES AND ACCES
897 2247 A 7761 1725 445 RPRRRĞTHEPSCVLQSFRVSAMMFPRVSTTLY RSAALLARTKNNIQRYGTINSVCSKENDKOR RSAALLARTKNNIQRYGTINSVCSKENDKOR RSAALLARTKNNIQRYGTINSVCSKENDKOR RSAALLARTKNNIQRYGTINSVCSKENDKOR RIESTKETISSEQDSKENTIKKELLGILGIKGMK VILISTUNVRTTKPPRRPILKSLEATLGULRRA TEYAPKKREPI-SPEL VAAASAVADSLPPDKQ TIKSELLSQLQQHEESRAQRDAKPKISFSNI ISDMKVARSATARVRSRPELRIQTDEGVTDNY QQEKTDOLKKRKINITGKRINITPDIMMAVTKE APETDISPSI WDVEFAKQLATVNEQPLQNG EBLIQWTKEGGL WEFTINNEAGFDDOSEFH EHITLEKHLESFPKQOPIRHYMELTYCOLSONO YLSVKQKVEHLEWRINYNFIKKDLIKESNIQP KLRPWKLFRNN TAKNOKOVEHLEWRINYNFIKKDLIKESNIQP KLRPWKLFRNN SQFGAVCKEPQEVVPGGGRSKRDPDLYQLI QRLPKSHSSLEGLIKALSQASTPPKESTSPEK RDMHDFFVGLMGKRSVQPDSTTDVNQENVP SGILKYPRRAE 859 2249 A 7785 179 703 PPHLGASSNTFRLQVOTQESKAQKEVKNOFE SSKRERQMANQLAWREFLEYFOTFGLA ABSLTAGAIKKKRAPATLVPIVPLSFLTYGYDL QSFGAVCKEPGLAWANGLERQLIM QSEMERQMANQLAWREFLEYFOTFGLA ALSLTAGAIKKKRAPATLVPIVPLSFLTYGYDL QYGTLLERNKGEAEDILETEKSKLQLPRGMIT ESSIKAKKEOSSRPEN FORMANGLARDLERGLIM SEMRERQMANQLAWREFLEYFOTFGLA ALSLTAGAIKKKRAPATLVPIVPLSFLTYGYDL QYGTLLERNKGEAEDILETEKSKLQLPRGMIT ESSIKAKKEOSSRPENT FESSIKAAKFOSSRPTIN FESSIKAKKOSSRPTNINTSSDVAALH KAMWKGVOBATIDULTKRNNAQRQQIKAAY LQETGRADELARDLANDLTSSDVAALH KAMWKGVOBATIDULTKRNNAQRQQIKAAY LQETGRADELARDLANDLTSTSODFRN ALLSLAKGDRSEDGVSEDLANDSDARALTY AQFDADELARAMKGLGTIDEDTLIELIASKIN KERDDINVYREELKRDLANDLTSTSTODFRN ALLSLAKGDRSEDGVSEDLANDSDARALTY AQFDADELARAMKROGTIDEDTLIELIASKIN KERDDINVYREELKRDLANDLTSTSTODFRN ALLSLAKGDRSEDGVSEDLANDSDARALTY AQFDADELARAMKROGTARRNY TKYSKHDMINKVLDLELKGDEKCLTAIVKGA GERKEGTDVNYNTILTTSTYPQLRRVFQKY TKYSKHDMINKVLDLELKGDEKCLTAIVKGA GERKEGTDVNYNTILTTSTYPQLRRVFQKY TKYSKHDMINKVLDLELKGDEKCLTAIVKGA GERKEGTDVNYNTILTTSTYPQLRRVFQKY TKYSKHDMINKVLDLELKGDEKCLTAIVKGA GERKEGTDVNYNTILTTSTYPQLARDLAND ALLSLAKGDRSCKEYNVLIQQLQKOT DLMQDTSRLLDPYRIQGLDVPKLRRHCEGR QFRFSERLRGLIGREFCTOTINATLGCVLHRL ADLEQRIPKAQDLERSGLANDLEKLQAMAP NILGERNNYCGARGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRRYGGT CHRKANSTISLASVAGAGTAATRRYGKGTAARRRY		ł	ł				
RPLSRIPILSGSPETSAAAMILITVRHGTYZY RSSALLARTKNINGRYFGTNSVICSKKDKOGSV RTEETSKETSEQDSEKENTKKDLIGIIKGMK VULSTVIVAVITKEPPKREPILKSLEATLGRIRAR TEYSAPKREIELISPELVAAASAVADSLPFDKQ TIKSELLSQLQQHEEESRAQRDAKRPKISFDKI ISDMKVARSATARVSSPELRIQPDEGYDNIVP GQEKTDDLKKRKNIFTGKRINFDMMAVTKE APETDTSPSJ.WDVPAKQLATVNEOPLQNGF EELIQWTKEGKL WEPFINNEAGFDDDGSEPH EELIQWTKEGKL WEFPINNEAGFDDDGSEPH VI.SVKQKVEHIEWFRNYTNEKKDILKESNIQP KIRPWKILFRIN VI.SVKQKVEHIEWFRNYTNEKKDILKESNIQP KIRPWKILFRIN ROMEN TO SCOTTOPPAGSCSTGTMRIMLIFTALLAFELA QSFGAVCKEPEGEVVPGGGRSKFUNTAKLOFELX QSFGAVCKEPEGEVVPGGGRSKFUNTAKLOFELX RDMHDFFVGLIMGKRSVQPDSPTDVNQENV SGILKYPPRAE 899 2249 A 7785 179 703 PFHLGASSNTFRQVQTQESKAQKEVKMGFI PKSMRESMRNGKEFMLNNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYPGTTFGLA AISLIAGAIKKKPATLVPIVLSFILTYQVDL GYGTLLERMGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMKGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMKGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMKGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMKGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFALVPIVLSFILTYQVDL GYGTLLERMKGEAEDILETEKSKLQLPRGMTI FESIEKARKGGAFASPSINGVLTQRQQKGAAY LQETGKFILDSTLKAATTGATARTQRQQKAAY LQETGKFILDSTLKAATTGATARTQRQQKAAY LQETGKFILDSTLKAATTGATARTQRQQKAAY LQETGKFILDSTLKAATTGATARTQRGATAATTNATGTI CHKRANTSTISLASVAGLTAAATRYTINPFGTT LEGYARVGGYTFTAAAAVFGLTTCISAVI GRANTYTIAAAATTAAAATTAAATTAAATTAAATTAAATTAA	807	2247	Α	2761	1225	445	
RSSALLARTKNNIQRYEGINSVICSKENGKOM RIESTSKETISEQDSEKENTKKELGILGIKGMK VELSTVAVRTTKPPRRPIKSLEATLGRLRRA TEYSAPKKREELPSEL VAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDARPPISSSNI ISDMKVARSATARVRSRPELRIQTDEGYDYNY GQEKTDOLKKRKINFTGKRINDFDMMAVTKE APETDISPSLWDVEPAKQLATVNEOPLONG EBLIQWTKEGKL WEFPINNEAGFDDGSEFF EBLIFLEKHLESFFKQOPIRHMELYTCGLSKNOP YLSVKQKVEHIEWFRNYPNEKKDLKESNOP YLSVKQKVEHIEWFRNYPNEKKDLKESNOP YLSVKQKVEHIEWFRNYPNEKKDLKESNOP KLRPWKFLFRNN KLRPWKFLFRNN SCOTTOPPAQSCSTGTMRIMLFTALLAFSLA QSFGAVCKEPQEEVVPGGGSKRDPDLYQLL QRLFKSHSLEGLLKALSQASTDPKESTSPL QRLFKSHSLEGLLKALSQASTDPKESTSPL QRLFKSHSSLEGLLKALSQASTDPKESTSPL RDMEDFFVGLMGKRSVQPDSTTDVNQENVP SFGILKYPPRAE RDMEDFFVGLMGKRSVQPDSTTDVNQENVP SFGILKYPPRAE FSKSMNESMKNQKEPMLMNARLQLERQLIM QSEMRERQMAMQLWSEFLKYGTTFGLA AISLTAGAIKKKRAPALVPIVPLSFLTTVYYDL GYGTLLERMKGEADEDLETEKSKLQLFRGMT FESIEKARKEGSRFTIDK QSEMRERQMAMQLWSEFLKYGTTFGLA AISLTAGAIKKKRAPALVPIVPLSFLTTVYYDL GYGTLLERMKGEADEDLETEKSKLQLFRGMT FESIEKARKEGSRFTIDK QSEMRERQMAMQLWSEFLKYGTTFGLA AISLTAGAIKKKRAPALVPIVPLSFLTTVYYDL GYGTLLERMKGEADEDLETEKSKLQLFRGMT FESIEKARKEGSRFTIDK QSEMRERQMAMQUSEFLKQAWFITFSLOGE GRDKDTFSKMAMVSEFLKQAWFIENEOBY VQTVKSSKGGPGSAVSPTFINFSSDLOFA QFGREDDTLKKALTGHLEEVVLALLKTT AQFDADELRAAMKGGTTEDTLELLASKTN KEIRDINEVYREELKRALTAKDITSDTSGDFRM ALSLAKGPRSEDFGVNEDLANDGRQKIKAY LQETGRFLDETIKKALTGHLEEVVLALLKTN REPRODNENDKAFYCKMYGTISHCKAR AUSTRAL GERKGTDVNVFNTILTTRSYPQLRRFGKY TYSKHSHDMINKLAGGTOEDTLELLASKTN KEIRDINDKAFYCKMYGTYMKALTRANG SKEMDNDIKAFYCKMYGTYMKALTRANG SKEMDNDIKAFYCKMYGTYMKALTRANG ANTON COMPANIA CHARLOST CHAR	057	4247	^	//01	1725	1 443	
RIEETSKETSESQDSEKENTIKEDLLGIIKGMEN VELSTVAVRTTEPFKRPLKSLEATGGILRRA TEYAPKKRIEPLSPEL VAAASAVADSL PPIGE TKSELLSOLQOHEESTA QRD AASAVADSL PPIGE PERIOTIS SUMDVERSTANDERS PLOT OR OKKTOL THE SELLSOLQOHEESTA QRD AKREPKISFS NI ISDMKVARSATARVRSRPELRIGFDEGYDNYP GOEKTDLKKRKNIFTGKRI NIFDMMAYTKE APETDTS'SL MVDYERAQLATINTEGPLQNGF EBELIQWTKEGKL WEPFINNEAGFDDDGSEFH EHITLEKHLESPFKOPIRHFMELTYCOLSKNP YLSVKGKVEHIE WERNYFNEKKDL KESNIOP KLRPWRLFRNN YLSVKGKVEHIE WERNYFNEKKDL KESNIOP KLRPWRLFRNN VLSVKGKVEHIE WERNYFNEKKDL KESNIOP KLRPWRLFRNN VLSVKGKVEHIE WERNYFNEKKDL KESNIOP KLRPWRLFRNN GSCOTTOPPAQSCSTGTMRIMLIFTALLAFSL QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGARSKRDPLLYQLA QSFGAVCKEPGEVVPGGGESKRDPLLYQLA QSFGAVCKEPGEVVPGGARSKRDPLLYQLA QSFGAVCKEPGEVVPGGARSKRDPLLYQLA QSFGASKGTDVARAFULPIVPLSFLTYQTVDL QSFGREQMAMQLAWSERFLLYQLAYTGTTFCAG QSFGATAGATLAFTLOYOTGESKAQCEFTEEFLFSSLQ QSFGATAGATLOYTCHETAGATLAFT QSFGATAGATLAFTLOYOTGESKAQCEFTEEFLFSSLQ QSFGATAGATAGATLAFTLOYOTGESKAQCEFTEEFLFSSLQ QSFGATAGATAGATLAFTLOYOTGESKAQCETTAGATCA QSFGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT	l	1				}	1
VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYPAKKREPLSPEL VAAASADSLPFDKQ TTKSELLSQLQQHEEFESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQPEDNOF GQEKTDLIKKRNIFTGRALNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNOF EELIQWTKEGKLWEFFINNEAGULATVNEQPLQNOF EELIQWTKEGKLWEFFINNEAGULKESNIQP YLSVKQKVEHEWFRNYTPKEKKDILKESNIQP KLRPWKFLFRNN 898 2248 A 7775 85 496 SCQTTQPPAGSCSTGTMRIMLIFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLRKSHSSLEGLILA ALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQFDSPTDVNQENVP SFGILKYPPRAE 899 2249 A 7785 179 703 PFHLGASSNITRILQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLLM QSEMRERQMAMQLAWSERLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQVDL GYGTLLERMKGGABDLLETEKSKLQLPRGMI QSEMRERQMAMQLAWSERLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQVDL GYGTLLERMKGGABDLLETEKSKLQLPRGMI FSSIEKARKEQSRFFIDK 900 2250 A 7789 1465 300 VWEPLKSYKIRSPSLHCQCEFREEFLFSSLQE GRONDTISKMAMVSEFLKQAWFIENEOGY VQTVXSSKGGFGAVSPYTFNSSDVAALH KAIMVKGVDEATIIDLTKRNAQAQQIKAAY LQETGREIDETLKKALTGHLEEVVLALLKTP AQFDADELRAMKGLGTDEDTLIELLASKTN KEIRDINRYYREELKRDLADSDARALYEA GERRKGTIDVNYFNTILTTRSYPQLRRVCKY TKYSKHDMNKVLDLEKGDETLIELLASKTN KEIRDINRYYREELKRDLADSDARALYEA GERRKGTIDVNYFNTILTTRSYPQLRRVCKY TKYSKHDMNKVLDLEKGDETLIELLASKTN KEIRDINRYYREELKRDLADSDARALYEA GERRKGTIDVNYFNTILTTRSYPQLRRVCKY TKYSKHDMNKVLDLEKGDTALDSDARALYEA GERRKGTIDVNYFNTILTTRSYPQLRRVCKY TKYSKHDMNKVLDLEKGDTALDSDARALYEA GERRKGTONYNTRITLTRSYPQLRRVCKY TKYSKHDMNKVAQULDTRSTDTAPTKAGGA SQPFTTTASDARAPGKGTAARRQKGTAARR QKGTAARRQKGTAARRQKGTAARRQKGTAARR QKGTAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGGTAARRQKGTAARRQKGTAARRQKGGTAA							1
TEYAPKKIEPI.SPEL VAAASAVADSLIPDKQ TIKSELLSQL.QOJGEEGSRAQDAKPKJESSIN ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKIIFTGRALINIEDMMAVTKE APETDTSPSL WDVEFAKQLATVNEQPI.QNGF EELLQWYKEGKL WEFPINNEAGFDDDGSEFH EHIPLEKHLESFRQOPRIHPMELVTCGLSKNP YLSVKQKVEHEWFRNYFNEKKDILKESNIQP KLRPWKFLFRIN 898 2248 A 7775 85 496 SCQTTOPPAGSCSTGTMRIMILIFTALLAFSLA QSFGAVCKEPGEVVPGGGRKRDPDLYQLL QRIFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFYGLMGKRSVQFDSPTDVNQENVP SIGILKYPPRAE 899 2249 A 7785 179 703 PFHLGASSNTFLLQVTQTGEKAQKWGFI FISKSMNESMKNQKEPMLMNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYRGTFFGLA AISLTAGAIKKKKAFLVPLYDLAG GYGTLLERMKGEABDILETEKSKLQLPRGMT FESIEKARLEGGSFFIDK 900 2250 A 7789 1465 300 VVLPIKSYKIRSPSLHCQCEIFREEFLFSSLQE GRADITTSKMAMVSEFLKQAWFIENEEQBY VVTVKSSKGGFGSAVSPYPTIPNSSDVAALH KAIMVKQVDEATIDILTKRNNAQRQUKAAY LQEITGKLDETLIKKALTGHLEGGY VVLPIKSSKGGFSAVSPYPTIPNSSDVAALH KAIMVKQVDEATIDILTKRNNAQRQUKKAY LQEITGKLDETLIKKALTGHLEGGY VVLPIKSSKGGFSAVSPYPTIPNSSDVAALH KAIMVKQVDEATIDILTKRNNAQRQUKKAY LQEITGKLDETLIKKALTGHLEGKCTATVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRMVS RSEDDMDBIKAFYQKMYGISLCQAILDETKGD 901 2251 A 7796 2 807 VEFFIPQRARAGRARFSMGVLLTGRITLLSILVL ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERKGTDVNVFNTILTTRSYPQLRRVYCKY TKYSKHDMNKVLDLELKGDIEKCLTATVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRMVS RSEDDMDBIKAFYQKMYGISLCQAILDETKGD VEKILVALCGGN VEKILVALCGGN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERKGTDVNVFNTILTTRSYPQLRRVYCKY TKYSKHDMNKVLDLELKGDIEKCLTATVKCA TSKPAFFAEKLHQAMKGVGTTRHKALIRMVS RSEDDMDBIKAFYQKMYGISLCQAILDETKGD VEKILVALCGGN TARARGKGTAARRQKGTAARR	Į.						l
TIKSELLSQLQQHEEESSAQRDAKRPKISSNN ISDMKVARSATAVRSREPLSTDEGYDNYP GQEKTDDLKKRKNIFTGKILNIFDMMAVTKE APETDTISPSL WDVEFAKQLATVNEQPLQNGF EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFPINNEAG EBLIQWTKEGKL WEFFINNEAG EBLIQWTKEGKL WEFFINNEAG EBLIQWTKEGKL WEFFINNEAG EBLIQWTKEGKL WEFFINNEAG EBLIQWTKEGKL WEFFINNEAG EBLIQWTKEGKL WEFFINNEAG EBLIGWTKA	i i						
ISDMKVARSATÄRVRSRPELRIQFDEGYDNYP GOEKTDILKKKRNIFTGKRLNIPDMMATYKE GEKTDILKKKRNIFTGKRLNIPDMMATYKE APETDTSPSL WDVEPAKQLATVNEQPLQNGF EELIQWTKEGKL WEFPINNEAGFIDDDSEFH EHIFLEKHLESPFKQGPRHFMEUVTGGLSKNP YLSVKQKVEHBEWFRNYFNEKKDILKESNIQF KLFPWKFLFRNN YLSVKQKVEHBEWFRNYFNEKKDILKESNIQF KLFPWKFLFRNN SCQTTQPPAGSCSTGTMRIMLITALIAFSLA QSFGAVCKEPGEVVPGGGRSKRDPDLYQLL QRLFKSHSLEGLLKALSQASTDPKESTSPEK RDMIDFFVGLMGKRSVQFDSTYDLYQLL QRLFKSHSLEGLLKALSQASTDPKESTSPEK RDMIDFFVGLMGKRSVQFDSTYDLYQLL QRLFKSHSLEGLLKALSQASTDPKESTSPEK RDMIDFFVGLMGKRSVQFDSTDVNQENVP SFGILKYPPRAE PSTILLAFSLA QSFGAVCKEPGEVVPGGGSKAPDPLYQLL QRLFKSHSLEGLLKALSQASTDPKESTSPEK RDMIDFFVGLMGKRSVQFDSTDVNQENVP SFGILKYPPRAE PSTILLAFSLA QSFGAVCKEPGEVVPGGGSKAPDLYQLL QSEMRERQMAMQLAWSREFLKYFGTFFGLA ALSLAGAKKKKPAFLVPIVPGELA ALSLAGAKKKKPAFLVPIVPTGLA GVGTTLLERMKGGAEDILETEKSKLQLPRGMT FESIEKARKGGSSFITDK QSEMRERQMAMQLAWSREFLKYFGTFFGLA ALSLAGAKKKFAFLVPIVPTGLA GVGTTLLERMKGGASDILETEKSKLQLPRGMT FESIEKARKGGSASSPITTNYSDVAALH KAMVKGVDEATIDILTRENNAQAQQUKAAY LQFGTGKLDETLIKKALTOHLEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIELLASSTN KERDINRYVREFLKRALTOHLEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIELLASSTN KERDINRYVSREFLKRDLAKDITSDTSGDFRN KERDINRYVSREFLKRDLAKDITSDTSGDFRN KERDINRYVDREFLKRDLAKDITSDTSGDFRN KERDINRYVDREFLKRDLAKDITSDTSGDFRN KERDINRYVDREFLKRDLAKDITSDTSGDFRN KERDINRYVDREFLKRDLAKDITSDTSGDFRN KERDINRYVDRAFYQKMYGISLQALDETKGD YEKILVALCGGN YEKILVALCGG							l
GĢEKTDDLKKRKNIFTGKRLNIFDMAVTKE APFDTISPS, WDVEPAKQLATVNEQPLQNGF EBLIQWTKEGKL WEPPINNEAGFDDDGSEFH EHIFLEKHLESPFKQOPIRHFMELVTCGLSKNF YLSVKQKVEHIBWFRNYFNEKKDILKESNIGF KLRPWKRLFRNN 898 2248 A 7775 85 496 SCQTTQPFAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEVVPGGGRSKRDPLYQLL QSFGAVCKEPQEVVPGGGRSKRDPLYQLQ QSFLKSHSSLEGLLKALSQASTDPKESTSPEK RDM:DFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE 899 2249 A 7785 179 703 PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAQIAWSEKLYVGTFTGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQVDL GYGTLLERMKGEAEDILETEKSKLQLPRGMT FSSIEKARKEQSRFFIDK 900 2250 A 7789 1465 300 VWEPLKSYKRISPSILHQCEIFREEFLFSSLQE GRONDITSKKAMMVSEFLAQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNSSDQAALH KAMVKGVVDEATIBLLTKRNNAQRQIKKAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQPDADELRAAMKGLGTDEDTLIELLASKTN KEIRDINRVYREELKRDLAKDITSDTSGDFIN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRAGTDVNVINTITTISYPQLRRVPQKY TKYSKHDMNKVLDLELKGDIEKCLTAVKCA TSKP4FFAEKHLHQAMKGVGTBKALVIKGD 901 2251 A 7796 2 807 VEFHPQRARAGAGAPSMGVLLTQRTLLSLVL ALLSFRAASGAGLOLVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINDVPKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGINFFHQH SVGRVFSKWGESPPRSKRRSPHQALRKGVGR TRASRKGKTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAAR							
APETDISPSI, WDVEPAKQLATVNEOPLONGEF EELIQWTKEGKI. WEPPINPAGEDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN 898 2248 A 7775 85 496 SCOTTOPPAQSCSTGTMRIMLIFTALIAFSLA QSFGAVCKEKPGEVVPGGRSKAPDPLVQLL QRIFKSHSSLEGLKALSQASTDPKESTSPEK RDMHDFFVGLMGRSVQPDSPTDVNQENVP SFGILKYPFRAE 899 2249 A 7785 179 703 PFHLGASSNTTRILQVQTQESKAQKKMGFI FSKSMMESKINQKEFMLMMARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGBAEDLETFKSKLQLPRGMT FSSIEKARKEQSRFTIDK 900 2250 A 7789 1465 300 VWFPLKSYKRSPSLHCQCEIFREEFLFSSLQE GRUKDTFSKMAMVSEFLKQAWFIENERQEY VQTVKSSKGOFGSAVSPFIDKSVALLKTP AQFDADELTAAMKGLGTDETMLEILASKTN KEIRDINRVYREELKRDLAKDTSDTSPRN ALLSLAKGOFGSAVSPFIPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELTAAMKGLGTDEDTLIEILASKTN KEIRDINRVYREELKRDLAKDTSDTSGDFRN ALLSLAKGGRSAFSFORDETMEDTLEILASKTN KEIRDINRVYREELKRDLAKDTSDDFRN ALLSLAKGGRSEDGVVEDLADSDARALYFA GERRKGTDVNVFNTILTTRSYPOLRRVFQKY TKYSKHDMNKVLDLELKGDEKCLTAVKCA TSKRAFFAEKLHQAMKGVGTRHKALIRINVS RSEIDMNDIKAFYQKMYGISLCQALDETKGD YEKILVALGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFFSMA MAAIGSCKEYNYLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSETLRGLGRRCFLQTLNATLGCVLHEL ADLEGRIFKAQDLERSGLTAGKGTAARRQ	[
EBLIQWTKEGKI.WEPPINNEAGFDDGSEFH EHHFLEKILESPFKQOPRIEMPLVTOGLSKINP YLSVKOKVEHIEWFRNYFNEKKDILKESNIOF KLRPWKFLFRNN 898 2248 A 7775 85 496 SCQTTOPPEQSCSTGTMRIMLLFTAILAFSLA QSSGAVCKEPQEVVPGGGRSKRDPLYQLL QRLFKSHSLEGILKALSQASTDPKESTSPEK RDM:IDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE 899 2249 A 7785 179 703 PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEPMLMMARQLQLERQLIM QSSEMRERQMAQIAWSKEFMLMMARQLQLERQLIM QSSEMRERQMAQIAWSKEFMLMMARQLQLPRGMIT FSSEKARKEQSRFFIDK 900 2250 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRENDTTSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNSSDVAALH KAMVKGVDEATIDLI.TKRNNAQRQQKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQPDADELRAAMKGLAGTDEDTLIELLASRTN KEERDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVNTNITLTSLYQLRRVPQKY TKYSKHDMNKVLDLELKGDIEKCLTAVKCA TSKFAFFAEKHLQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YVEKLVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGAGRAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLQCJLQKQT DLMQDTSRLLDPYRIQCIQLOKQT DLMQDTSRLLDPYRIQCIQLOKQT DLMQDTSRLLDPYRIQCIQLOKQT DLMQDTSRLLDPYRIQCIQLOKQT DLMQDTSRLLDPYRIQCIQLOKQT NILGLRNNIYCMAQILDNSDTAEPIKAGRGA SQPPTPTPASDAFQKLEGCGR-HGYHFPMH SVGRVFSKWGESPPRSKRISPHQALRKGVRR TRPSRKGKRLMTRQQLPR 1RPSRKGKRLMTRQQLPR 1RPSRKGKRLMTRQQLPR 1RPSRKGKRLMTRQQLPR 1RPSRKGKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRPQKGTAARRRQKGTAARRRQKGTAARRPQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRPQKGTAARRRQKGTAARRPQKGTAARR							
BHIFLEKHLESPRKQGPRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKDILKESNIQF KLRPWKFLFRNN 898 2248 A 7775 85 496 SCQITTQPPAQSCSTGTMRIMLIFALLASSLA QSFGAVCKEPQEEVVFGGRSKRDPDLYQLL QRLFKSHSSLEGLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE 899 2249 A 7785 179 703 PFHIGASSNITRLQVQTQESKAQKEVKMGFI FSKSMESMENQKEPMMNNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYFGTFFGLA ASLTAGAKKKKPAFTMMNNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYFGTFFGLA ASLTAGAKKKKPAFTMMNNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYFGTFFGLA ASLTAGAKKKKPAFTMLNNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYFGTFFGLA ASLTAGAKKKKPAFTMLNNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYGTFFGLA ASLTAGAKKKKPAFTMLNNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYGTFFGLA GRENKOTTSKMAMAVSEFLKQAWFIENEQEY VQTVKSSKGGPGSAVSPYTPNSSDVAPY VQTVKSSKGGPGSAVSPYTPNSSDVAPY VQTVKSSKGGPGSAVSPYTPNSSDVAPY LQETGKPLDETLKKALTGHLEEVLALLKTP AQPTADELRAMKGLGTDEDTLIELLASKTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGORSEDFGVNEDLANSDARALVFA GERKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMIKVLDLELKGDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YERILVALCGGN VEFHPORARAGARAPSMGVLITARTLGCVLHRL ADLEQRLPFAGDLERSGAMEDLEKLQMAPP NILGRNNIYCMAQLLDNSDTAGTLGCVLHRL ADLEQRLPFAGDLERSGAMEDLEKLQMAPP NILGRNNIYCMAQLLDNSDTAGTHAGCVLHRL ADLEQRLPFAGDAFTKAGRGA SQPFITTPASDAFQRKLGCRFLHGYYRFMH SVGRVFSKWGESPINSRRISPHQALRKGVRR TRPSRKGKTMARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRRQKGTA							
Sykokyvehiewprnyfnekkdilkesniof Klrpwkfifnnn							
SUMPRICE STATE S						:	
898 2248 A 7775 85 496 SCOTTOPPAGSCSTGTMRILITAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQL QRLFKSHSSLEGLLKALSQASTDFKESTSPEK RDMIDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE 899 2249 A 7785 179 703 PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMIRNQKEFMLMNARQLERQLIM QSEMRERQMAMQLAWSREFLKYFGTFGLA AISLTAGAIKKKRPAFLVPIVPLSFILTYQYDL GYGTLERMKGEAEDLERKSKLQLPRGMIT FESIEKARKEQSRFFIDK 900 2250 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRIKDTTSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTPNSDVAALH KAIMVKQDEATIDILTERKSKLQLPRGMIT FESIEKARKEQSRFFIDK 901 2251 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRIKDTISKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTPNSDVAALH KAIMVKQDEATIDILTERKSKLQLPRGMIT FESIEKARKEQSRFFIDK 902 2250 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRIKDTISKMAMVSEPLKAGVAFTENDAGVAAY LQETGKPLDETILKKALTOHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTILEILASRTIN KEIRDINKVYREELKRDLAKDITSDSDARALYEA GERKKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMINKVIDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS REIDINMINKVIDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS REIDINMINKVALDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS REIDINMINKVALDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS REIDINMINKVALDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS REIDINMINKVALDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS REIDINMINKVALDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS REIDINMINKVALDLEIKDDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTLAILETKGD YERL VALCEGR PAGFTENDAG TAIRT	,						
PST AVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDFKESTSPEK RDM-IDP-FVGLMGRVSQPDSPTDVNQENVP SFGILKYPPRAE 899 2249 A 7785 179 703 PFFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFILKYGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIBKARLEQSRFFIDK 900 2250 A 7789 1465 300 VVLPLKSYKIRSPSILHQQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDLITKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKITP AQFDADELRAAMKGLGTDEDTLIELLASRTN KEIRDINRYVREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMMKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLIQAMKGVGTRIKALRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEHPPQRARAGRAPSMGVLLTQRTILSLVL ALLFPSMASMAAIGSCSKEVRVLLGQLQKQT DLMQDTSRILDPYRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTILNATLGCVHRIL ADLEQRLPKAQDLERSGLNEDLEKLQMARP NIGLERNNYCMAQULDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPMRSRRHSPHQALRKGVRR TRFSRKGKILMTRGQLDPK TRFSRKGKILMTRGQLPK OKGTAARRRQKGTAARRRQKGTAARRRQK GTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQK TAARRRQKGTAARRRQKGTAARRRQKGTAARRRQK TAARRRQKGTAARRRQKGTAARRRQKGTAARRRQK TAARRRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARR	898	2248	Α	7775	85	496	
899 2249 A 7785 179 703 PFHLGASSNITRLQVQTQESKAQKEVKMGFI FSKSIMSKENKOKEFMLMNARLQLERQLIM QSEMRERQMAMQLAWSREFLKYFGTFFGLA AISLTAGAIKKKRAPTLVPIVLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK 900 2250 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREFLFSSLQE GRDKDITSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGFGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDLITKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKITP AQFDADELRAAMKGLGTDEDTLIELLASRTIN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTIDVNVNTTILTTRSYPQLRRVFQKY TKYSKHDMINKVLDLELKGDIEKCLTANVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYNVLLGQLQKQT DLMQDTSRLLDPYRIQGLDVFKLREHCREP GAPPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTFTBASDAFQRKLEGCRFLHGYHRFMH SVGRVSKWGESPNSRRHSPHQALRKGVRR TRPSRKGKRLMTRQQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAARRQKG TAARRAGAGSULRAMVTUNPPGTF LEGVAKVGQYTFTAAAAGAVFGLTTCISAHV REKPIDDFLNYFT.GGCAGGLTLGASKTHNYGIG CHRKAYSTTSIASVAGLTAAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAAGAVGLTTCISAHV					33	1,50	
899 2249 A 7785 179 703 PFHLGASSNITRLQVQTQESKAQKEVKMGFI FSKSMKNOKEFMLMNARLQLERQLIM QSEMBERQMAMQWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETKSKLQLFRGMTI FESIEKARKEQSRFIDK 900 2250 A 7789 1465 300 VVLPLKSYKIRSFSLHCQCEIFREFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEGQEY VQTVKSSKGGPGSAVSPYPTFNSSDVAALH KAIMVKGVDEATIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTIN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSGOVSPOPTFNSSDVAALH KAIMVKGVDEATIDILTTKRNNAQRQVIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTIN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSGOVNEDLADSDARALVEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMKVUDLELKGDIEKCLTATVKCA TSKPAFFAEKLIPGOVNEDLADSDARALVEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMKVUDLEKGDIEKCLTATVKCA TSKPAFFAEKLIPGDVSCHALLKTP ALLSPSMASMAGIGSCSKESKYVLLGQLQKQT DLMQDTSRLLDPVIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRIPKAQDLERSGLNIEDLEKLQMARP NILGERNNIVCMAQLLDNSDTAEPTKAGRGA SQPPTTPASDAFQKLEGCGRFLHGVHRFMH SVGRVFSKWGESPNSRRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARTGAAGATHNYGIG							
899 2249 A 7785 179 703 PFHLGASSNITRLQVQTQESKAQKEVKMGFI PSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMERRQMAMQLAWSRETLK YPGITFGLA AISLTAGAIKKKRPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK 900 2250 A 7789 1465 300 VVLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEGQEY VQTVLKSSKGGPGSAVSSYPTIPNSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITISDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALRIMVS RSEIDMNDIKAFYQKNYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCEVERVLLQQLQKQT DLMQDTSRLLDPYTRIQGLDVFKLEEHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLFRL ADLEQRIPKAQDLERSGLNIEDLEKLQMARP NILGIRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTTPASDAFQRKLEGCKFLHGYHRFMH SVGRVPSKWGESPNSRRHSPHQALRKGVRR TRPSRKGKRLMTRQQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAARRTHNYGIG CHR	j						
899 2249 A 7785 179 703 PFHLGASSNITFLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMMARLQLERQLIM QSEMBERQMAMQWSREFLKYKTGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTILERMKGEAEDILETKSKLQLFRGMT FESIEKARKEQSRFIDK 900 2250 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREFLFSSLQE GRDKDITFSKMAMVSEFLKQAWFIENEGDEY YQTVYKSSKGGPGSAVSPYPTFNYSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTF AQFDADELRASMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNEVLDLELKGDEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKNYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCREPP GAFPSEFTLRGLGRRCFLGTLNATLGCVLHRL ADLEQRLFKAQDLERSGLNIEDLEKLQMARP NILGLRNNIVCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQKLGCFFLHGYHRMH SVGRVFSKWGESPNRSRRHSPHQAHRGVRR TRPSRKGKRLMTRQQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAARRRQKGTAARRQ		i					
FSKSMNESMKNQKEFMLMNARIQLERQLIM QSEMRERQMAMQIAWSREFILKYFGTFFGLA AISLTAGAIKKKPAFLVPIVPSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKAAKEQSRFTDIN 900 2250 A 7789 1465 300 VWIPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGFGSAVSPYFTFNPSSDVAALH KAIMVKGVDEATIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIELLASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTATVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLGGNPCFLQTLNATLGCVLHRL ADLEQRLPKAGGGASKEYRVLLGQLQKGT DLMQDTSRLDDYRRQGLDVPKLREHCRERP GAFPSEETLRGLGRCFCLQTLNATLGCVLHRL ADLEQRLPKAGGESPNSRRHSPHQALRKGVRR TRPSRKGKRLMTRQQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAARRRQ	899	2249	A	7785	179	703	
900 2250 A 7789 1465 300 VVILEKSYKRSPSLHCQCEIFREFLFSELQE GRDKDTSSKMAMVSEFLKQAWFIENEGEFY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQIKAAY LQETGKFLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDAWFIENEGEFY TKYSKHDMNKVLDLELKGDUEKCLTAIVKCA GERKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDUEKCLTAIVKCA TSKPAFFAEKHQAMKGVGTHAKALRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGRAPSMGVILTQRTILSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYRINGGLDVPKLREHCRERP GAFPSEFILRGIGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSCLNEDLEKLGWARP NILGLRNNIYCMAQLLDNSDTAEPIKAGRGA SQPPTPTASDAFQRKLEGCRFLHGVHRFMH SVGRVPSKWGESPNRSRHSPHQALRKGVRR TRPSRKGKLIMTRGQLPR 902 2252 A 7802 2 721 TAARRQKGTAARRQKGTAARRRQKGTAGAGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG				,,,,,			FSKSMNESMKNOKEFMI MNARI OF FROLIM
900 2250 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREFLFSSLQE GRDKDTFSKMAMVSEFLKQA WFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQIKAAY LQETGKPLDEATLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KERDINRYVREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVVEDLADSDARAL YEA GERKGTDVNVFNTILTTRSYPQLRVFVGKY TKYSKHDMNKVLDLELKGDLEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKNYGISLCQAILDETKGD YYEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLSPSMASMAAIGSCSKEVRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCREPP GAFPSETLIRGLGRRCFLQTILNATLGCVLHRL ADLEQRIPKAQDLERSGLNIEDLEKLQMARP NILGLRNINYCMAQLLDNSDTAEPTKAGRGA SQPFIFTASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTA							
900 2250 A 7789 1465 300 VWLPLKSYKIRSPSILHCQCEIPREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAMVKGVDEATIBILI TKRNAQRQCIKAAY LQETGKPLDETILKALTRNAQRQCIKAAY LQETGKPLDETLKKALTOHLEEVVLALLKTP AQPDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRYVEREILKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFTAEKLHQAMKGYGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEEHPQRARAGARAPSMGVLLTQRTILSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYRIQGLDVPKLREHCRERP GAFPSEETLRJGGRRCFLGTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLDNSDTAEPTKAGRGA SQPPTPTBADAFQRKLEGCRFLHGYHFPMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKG TAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRQK					i		
900 2250 A 7789 1465 300 VWLPLKSYKIRSPILHOQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKROLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQALLDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRILDPYIRIQGLDVFKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLFKAQDLERSGLNIEDLEKLQMARP NILGIRNNYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRISPHQALRKGVRR TRPSRKGKLHTRGQLPR TRPSRKGKLHTRGQLPR TAARRRQKGTAARRPQKGTAARRRQKGTAA RRQKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRR AGGVRGAGSRLRAMAFKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVQQYTFTAAAVGAVFGLTTCISAHV REKYDDPLNYFLGGCAGGLTLGARTHNYGIG	ŀ						
900 2250 A 7789 1465 300 VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQIKAAY LQETGKPLDETLIKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASKTIN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKLVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFFSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGIRNNYCMAQLLDNSDTAEPTKAGRGA SQPTTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRAGAVBAVTLNPPGIF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKKPDDPLNYFLGGCAGGLTLGARTHNYGIG							
GRDKDTFSKMAMVSEFLKQA WFIENEGE Y VQTVKSSKGGGSAVSPYPTFNPSSDVAALH KAIMVK GVDEATIIDIL TKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNIEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGRAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVFKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGRINNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAAARRRQKG TAARRRQKGTAARRRQKGTAAARRRQKG TAARRRQKGTAARRRQKGTAAARRRQKG TAARRRQKGTAARRRQKGTAAARRRQKG TAARRRQKGTAARRRQKGTAAARRQKG TAARRRQKGTAARRRQKGTAAARRQKG TAARRRQKGTAARRRQKGTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLITCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG	900	2250	Α	7789	1465	300	
VQTVKSSKGGPGSAVSPYPŤFNPSSDVAÁLH KAIMYKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIELASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKLVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRIPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRPQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARTATATATATATATATATATATATATATATATATAT		1	ſ	. [
KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLIDETLKKALTGHLEEVVLALLKTIP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHIDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGYGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGRAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYRIQQLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTBASDAFQRKLEGCRFLHGYHFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRAGAGVGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGATTHNYGTG		İ					
LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAPYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVPSKWGESPNRSRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAARRQ QKGTAARRRQKGTAARRQKGTAARRQ GKGTAARRRQKGTAARRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRR QKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTICISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGTG)	}	- 1				KAIMVKGVDEATIIDILTKRNNAOROOIKAAY
AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAARRR QKGTAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRQVGTAARRRQKG TAARRRQKGTAARRQVGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRQVGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRQVGTAARRRQKG TAARRRQKGTAARRQVGTAARRRQKG TAARRRQKGTAARRQVGTAARRRQKG TAARRRQKGTAARRQVGTAARRRQKGTAGTATANYGTIGGGCAGGGTTLGARTTHNYGIGG	1	l					
KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTFTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRAGKGTAARRAGKGTAARRAGKGTAARRAGKGTAARRAGKGTAARRAGKGTAARRAGKGTAARRAGKGTAARR			Ì	1			AQFDADELRAAMKGLGTDEDTLIEILASRTN
GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRRQKGTAARRRQKG GTAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRQKGTAARRRQKG TAARRRQKGTAARRQKGTAARRRQKG TAARRRQKGTAARRQKGTAARRRQKGTAARRQKG TAARRRQKGTAARRQKGTAARRRQKGTAARRQKG TAARRRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRAGATAGATTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG				1			
TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLIDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRQKGTAARRLQKGTAARRRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRQKGT		ļ			1		ALLSLAKGDRSEDFGVNEDLADSDARALYEA
TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEFTLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRLQKGTAARRRQKG RRRQKGTAARRPQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG	[[ſ	ĺ	i		GERRKGTDVNVFNTILTTRSYPQLRRVFQKY
RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN 901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRRQKGTAARRRQKG RRRQKGTAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG		İ	1				TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA
901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRRQKGTAARRRQKG RRRQKGTAARRPQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG		ļ	j	ļ	J		TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS
901 2251 A 7796 2 807 VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRLQKGTAARRRQKGTAARRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRQKGTAARRQKGTAARRQKGTAARRRQKGTAARRAGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG		ľ		1	1		RSEIDMNDIKAFYQKMYGISLCQAILDETKGD
ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRLQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRQKG TAARRRQKGTAARRQKGTAARRQKGTAARRQKG TAGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG							
DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRLQKGTAARRRQKGTAARRAGATGTTDGTDCATAGATARTATATRATATATATATATATATATATATATAT	901	2251	A	7796	2	807	
GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRLQKGTAARRRQKGTAAAYRAVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG	i		- 1		ĺ		
ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAA RRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGTAARRRQKG TAARRRQKGTAARRRQKG TAARRRQKGTAARRQKGTAARRQKG TAARRRQKGTAARRQKGTAARRQKG TAGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG		ľ	i		1		
NILGERNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRLQKGTAARRRQKGTAARRGQKGTAARRGQKGTAARRRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGTAARRQKGT	ľ	ľ	ł	İ	l	ł	
SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR 902 2252 A 7802 2 721 TAARRRQKGTAARRLQKGTAARRRQKGTDCHRKAYSTTSIASVAGLTAAAYRVTLNPPGTFLEGVAKVGQYTFTAAAVGAVFGLTTCISAHVREKPDDPLNYFLGGCAGGLTLGARTHNYGIG		ļ			l	ł	
902 2252 A 7802 2 721 TAARRRQKGTAARRATAARRQKGTAARRATAARRQKGTAARRATAARRQKGTAARRATAARRQKGTAARRATAARRQKGTAARRATAARRQKGTAARRATAARATATAARTATAARRQKGTAARRATAARATAARATAARATAARATAARATAARATA	ł	- 1					
902 2252 A 7802 2 721 TAARRRQKGTAARRQKGTAGRTDCHRKAYSTTSIASVAGLTAAAYRVTLNPPGTFLEGVAKVGQYTFTAAAVGAVFGLTTCISAHVRGKGTARTHNYGIG	ĺ	- 1	- 1		. [ľ	
902 2252 A 7802 2 721 TAARRRQKGTAARRARRQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARGKGTAARRATARARGKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARQKGTAARRARARARARARGKGTAARRARARAR	i		1		ĺ		
RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG	002	2262	_	7002			
QKGTAARRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG	902	2252	A	/802	4	721	
TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG	1		. 1	-	ł	ľ	
AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG		1		1			_*
CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG		}	- 1				
LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG	1	ł	- 1	- 1		i	
REKPDDPLNYFLGGCAGGLTLGARTHNYGIG		ļ	1	.			
		ļ		1		ł	-
	ĺ	1	- 1	1	.[1	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1							AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN 09/496	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	delice		914	correspondi ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uciicc			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	beque	/=possible nucleotide deletion, \=possible
1		1	ł	sequence	ł	nucleotide insertion
903	2253	Α	7807	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA
1						VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
1		1	l	1	1	RGDSPTDSQKDMIEIPLPPWQERTDESIETKR
1						ARLLYESRKRGMLENCILLSLFAKEHLQHMT
						EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
		1		Ī	ļ	EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
904	2254	A	7813	40	821	LFEKPR
304	2234	^	/613	40	021	GAGRALGHLETGAGDVAAALPARKFPRSLLG AGARLTGWTMNVFRILGDLSHLLAMILLLGK
		1				IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF
		1			j	ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF
ĺ		ĺ				DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL
'			ļ			WTFSIYLESVAILPQLFMISKTGEAETITTHYL
			1			FFLGLYRALYLANWIRRYQTENFYDQIAVVS
1.			ĺ	ĺ		GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL
		<u> </u>				RSYSSI
905	2255	Α	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA
						QTEMVRTLERKLEAKMIKEESDYHDLESVVQ
)			QVEQNLELMTKRAVKAENHVVKLKQEISLL
						QAQVSNFQRENEALRCGQGASLTVVKQNAD VALONLRVVMNSAOASIEOLVSGAETLNLVA
						EILKSIDRISEVKDEEEDS
906	2256	A	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL
		[:]			1.02	LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG
					'	HQPQTGSGESSGASGDKDHLYSTVCKPRSPK
1						PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA
		ļ.				TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK
						RPSLPSSPSPGLPKASATSATLELDRLMASLSD
						FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG
1						KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
1 1						NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
						GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG
						FHEREGRPYCRRDFLOLFAPRCOGCOGPILDN
1 1						YISALSALWHPDCFVCRECFAPFSGGSFFEHE
						GRPLCENHFHARRGSLCATCGLPVTGRCVSA
						LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY
						CQPCFLKLFG
907	2257	Α	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH
1000						YCKSQAWG
908	2258	A	7842	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
	į	ļ				SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET
			.			YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI
	ŀ					NMNLDVSDEKMAMFSLACILPFKYPAVLPEI
					,	TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV
						CILNATEWVREHASGYVSRDTSSSPTTGSTVQ
	1	1		j	ĺ	SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL
				Ì	Į.	SLSGFSMPGKPGVVCVEGPQSACEEFWARLR
				İ		KLNWKRILIRHREDIPFDGTNDETERQRKFSIF
1 1		1			l	EEKVFSVNGARGNHMDFGQLYQFLNTKGCG
1000	2052			2065		DVFQMFLWV
909	2259	A	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV
910	2260	_	7004	212	1071	LISSEILLIPSKYLFESK
טופ	2260	A	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHPVNARPGCVVTNACASGPCPPHANCPDI
				.	·	PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG
						SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD
	ļ	ļ	j		ļ	QQCPRGWWGSPTCGPCNCDVHKGFDPNCNK
		1				CALL STATE S

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV TASGCRVLYDACPKSLRSGVWWPQTKFGVL ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFRELSLLLDGLEINKTALDTME AKKLAQRLREVTGHTDHYFSQDVRVTARLL AHLLAFESHQQGFGLTATQDAHFNENLLWA GSALLAPETGDLWAALGQRAPGGSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS VVPPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ AERGARLPQNPVMNSPVVSVAVFHGRNFLR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGIHANVAAA LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT
911	2261	A	7890	21	806	EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE DKTFFGAFLL
. 912	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ PVPYILKKIFQDREAAATTGVSRDLCYVKELG VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL EILVKEDRDSGVNFQPEDTCARLRCSLHASLL VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH RHQLFINFRDLGWHK WIIAPKGFMANYCHGE CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV CIPTKLSPISMLYQDNNDNVILRHYEDMVVD
913	2263	A	7892	15	849	ECGCG ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS VGATCCYLLSSIFGKQLVVSYFPDKVALLQR KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ KHLQLNETSTANHIHSRKDT
914	2264	Α	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN RLSNITRPFFSKKKKKILV
915	2265	A	7909	3	641	HASGPGGLLRRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDQELKLIGEYGLRNKREV WRVKFTLAKIRKAARELLTLDEKDPRRLFEG NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS HVKRKNASKGQGGAGARDDEEEE
916	2266	A	7914	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC QVLILKHTHASLSLPSCQECFPSSIPSASHMVS HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS SILPTGDSWGMLACLCTVLWHLPAVPALNRT GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN YLGPPFNEPDFNPPRLGAETLPRATVDLEVW RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA TAELRRSLAHFCTSLQGLLGSIAGVMAALGY PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL LKELQTWLWRSAKDFNRLKKKMQPPAAAVT LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED THHQIATLTVLVAPENPVVEVREQAVEGGEV ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ ENGKVWSVASTVRFRVDRKDDGGIIICEAQN QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL PERAEAVGETLTLPGLVSADNGTYTCEASNK HGHARALYVLVVYGESRLRPTEGGGAPDP GAVVEAQTSVPYAIVGGILALLVFLIICVLVG MVWCSVRQKGSYLTHEASGLDEQGEAREAF LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSSVSSSLSPSAVVMACRWSTK ESPRWRSALLLLFLAGVYGNGALAEHSENVH ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK INCSWFIRANPGEIITISFQDFDIQGSRRCNLD WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR CGNGKCIPEAWKCNNMDECGDRSDEEICAKE ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	1	peptide		/=possible nucleotide deletion, \=possible
ļ		<u> </u>		sequence		nucleotide insertion
		j				VILRFTDFKLDGTGYGDYVKIYDGLEENPHK
						LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV
1		İ		•		NAARGFNATYQVDGFCLPWEIPCGGNWGCY
			i			TEQQRCDGYWHCPNGRDETNCTMCQKEEFP
						CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF
						CQPGNFHCKNNRCVFESWVCDSQDDCGDGS
1		1				DEENCPVIVPTRVITAAVIGSLICGLLLVIALG
		i				CTCKLYSLRMFERRSFETQLSRVEAELLRREA
		i				PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL
1						RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA
1		!				RSRHSGSLALVSADGDEVVPSQSTSREPERNH
i		1				THRSLFSVESDDTDTENERRDMAGASGGVAA
l.		ŀ				PLPQKVPPTTAVEATVGACASSSTQSTRGGH
1		į				ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ
j .		l				GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG
						REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL
		1				ASDQGQGLRQPYNATNPGVRPSNRDGPCERC
919	2269	Ā	7951	1674	1839	GIVHTAQIPDTCLEVTLKNETSDDEALLLC VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
319	2209	^	/931	10/4	1639	GGWNDVACHTTMYFMCEFDKKNM
920	2270	Α	7953	47	572	GGRASWPEQAKEPRREGHTDKQOTEDVLAA
920	2270	^	1933	47	3/2	GLRCLPHLPAICARRMSPAFRAMDVEPRAKG
		1				VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE
		l				HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW
			1			RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS
1			1			FSSSWPSAQHLTPSVFNPW
921	2271	A	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLONEF
		*		V-	012	MMVYSFRALSFKESTWATFQHGGEATKSRSL
1		Ì	, ·			SSTO
922	2272	A	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT
i i						VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN
		· ·	1			WELVKPN
923	2273	Α	7981	1	3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPPLL
						LLPLLLLPAGCRALEETLMDTKWVTSELAWT
1 1						SHPESGWEEVSGYDEAMNPIRTYQVCNVRES
1						SQNNWLRTGFIWRRDVQRVYVELKFTVRDC
1						NSIPNIPGSCKETFNLFYYEADSDVASASSPFW
						MENPYVKVDTIAPDESFSRLDAGRVNTKVRS
]				:		FGPLSKAGFYLAFQDQGACMSLISVRAFYKK
						CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN
1 1						AVEVSVPLKLYCNGDGEWMVPVGACTCATG
1						HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP
J J						NSRTTSPAASICTCHNNFYRADSDSADSACTT
		i			·	VPSPPRGVISNVNETSLILEWSEPRDLGVRDD
	.				1	LLYNVICKKCHGAGGASACSRCDDNVEFVPR
						QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS
						GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS
						SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG
1					l	IASTVTSQMNSVQLDGLRPDARYVVQVRART
					ļ	VAGYGQYSRPAEFETTSERGSGAQQLQEQLP
					ı	LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS
						DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA
	•			ĺ	ĺ	VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL
				l	l	KQPGRREVFVAIKTLKVGYTERQRRDFLSEA
1				ļ	l	SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME
i 1				1		NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM
]	1		
				1		KYLSEMNYVHRDLAARNILVNSNLVCKVSDF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	L Amino poid popular (AmAlonias CarCardia
NO: of			ID NO:			Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	•	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
June])	1 211	amino acid		
İ		1	1		of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		i	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	i	peptide		/=possible nucleotide deletion, \=possible
1	ļ			sequence		nucleotide insertion
		-			 	WDMSNQDVINAVEQDYRLPPPMDCPTALHQ
1	ĺ	1	1	1	ĺ	I A DOWN DOWN DOWN DOWN DOWN DOWN
		1	1	ł		LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA
1		1	l .	ł		SLKVIASAQSGMSQPLLDRTVPDYTTFTTVGD
j	Į.	1		ĺ		WLDAIKMGRYKESFVSAGFASFDLVAQMTA
1	1			1	1	EDLLRIGVTLAGHOKKILSSIODMRLOMNOT
i	ľ	1	ſ	ľ	İ	LPVOV
924	2274		7985		602	
924	22/4	Α	/985	1	-503	FRPRTKKATAMYLEHYLDSIENLPCELQRNF
			1			QLMRELDQRTEDKKAEIDILAAEYISTVKTLS
	1	l		i		PDQRVERLQKIQNAYSKCKEYSDDKVQLAM
1	l	1	J		}	QTYEMVDKHIRRLDADLARFEADLKDKMEG
	ł				1	
	l		ļ			SDFESSGGRGLKKGRGQKEKRGSRGRGRRTS
L	L		L			EEDTPKKKKHKGG
925	2275	Α	7994	447	589	LPCSFCAQCMSSFERVWLQQSHFHNPRWNSR
J .]	1]]	SPIRCYCQHWPHCVHC
926	2276	A	7996	925	582	GPCKVCCITLAIMLQCHSFYRKDVQVEHPKS
120	2270	l ^^	1770	123	302	OF OWACHTERING COST I KYDA A FALL COST I KYDA A FAL
			1			LNPKYSQIENFLSADMALKRKCLLSISDLDFW
1						IWDAQPVGIMQTLQNLKKIPNPGCFWSQAFQI
		ļ				RDTQPILPLGGRYYITIRQ
927	2277	A	7998	2	353	RIQRPLNSRSPNHSLFVKAELTAKQATMKLSV
	527.	1 **	,,,,	-	333	
ļ			1 :			CLLLVTLALCCYQANAEFCPALVSELLDFFFI
			1			SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ
						MSLQKRSLIAEVLVKILKKCSV
928	2278	A	8004	130	588	LAPLRCOPGTRTOPRSHPAANDPSAAMSAAG
			1			ARGLRATYHRLLDKVELMLPEKLRPLYNHPA
		i	1			
i			1			GPRTVFFWAPIMKWGLVCAGLADMARPAEK
						LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA
j] 1		ļ	VNFFVGAAGASQLFRIWRYNQELKAKAHK
929	2279	A	8007	2	1016	EFARRVFIAAREMSLLRSLRVFLVARTGSYP
				_		AGSLLRQSPQPRHTFYAGPRLSASASSKELLM
l l						
			1			KLRRKTGYSFVNCKKALETCGGDLKQAEIWL
						HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE
1 1			i i			GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL
						GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA
, ,						GPDREGSLKDQLALAIGKLGENMILKRAAWV
1						
i i						KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG
, ,			, ,			ALVICETSEQKTNLEDVGRRLGQHVVGMAPL
ļ <u>ļ</u>						SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ
						YVQPQGVSVVDFVRFECGEGEEAAETE
930	2280	A	8008	3	1679	NSRVWGPWTEPSAGSLRPMARKQNRNSKEL
	240V	A	0000		10/7	
, ,			,)			GLVPLTDDTSHAGPPGPGRALLECDHLRSGV
	- 1					PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN
	i				'	LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL
	l	1				TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK
1						
I	J					HTLLANNGFAISAALLMACSLQAGAFEMLIV
[[[[ĺ	ľ	GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG
l						QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF
						GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA
				l	İ	RAVKAFQTFLGKADVSQEVEEVLAESRVQRS
· .			}	ſ		
1						IRLVSVLELLRAPYVRWQVVTVIVTMACYQL
ı		ĺ		1		CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG
	ł		Į Į			IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF
	i					GTLTITLTLQDHAPWVPYLSIVGILAHASFCSG
					ļ	PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN
	į		ı			TANDERILLUTER KINGKPAAPHAGI VNWI SN
ļ	j				1	DALLOT I EDITOROT DETERMINE IN TOTAL TOTAL
						FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL
						FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL
						FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL , YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI
931	2281	A	8009	861	300	FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL , YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP
931	2281	A	8009	861	300	FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL
931	2281	A	8009	861	300	FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL , YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide .	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Laucine,
seq-	uence	j	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 .	}	1	1	peptide	'	/=possible nucleotide deletion. \=possible
1		ŀ	1	sequence		nucleotide insertion
						PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
		Ì				LIDYVRYMVENHGEDYKAMARDEKNYYQD
ĺ		i				TPKQIRSKINVYKRFYPAEWODFLDSLOKRK
						MEVE
932	2282	Ā	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
]	1.0022		} `	DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
		ĺ	1			QGRGQIPIPCPWPPPPPPPPPPGSPGPGCROFHO
						SLEAKARHPASVREMRGKVKMRRALRRAPA
1			i			STRASSROPNPK
933	2283	A	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
733	2203	Α.	8012	147	10//	
		l				ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
		l	1			NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
		1	<u>[</u>			RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
						PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
		1				NASQLITQRAQVSLLIRRELTERAKDFSLILDD
						VAITELSFSREYTAAVEAKQVAQQEAQRAQF
		ł				LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
1	l	i	l I	1		SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
						DNLVLNLQDESFTRGSDSLIKGKK
934 -	2284	Α	8023	255.	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
		ľ			i	RLRKFRELHLMRNEARKLNHQEVVEEDKRL
1			İ			KLPANWEAKKARLEWELKEEEKKKECAARG
		ĺ				EDYEKVKLLEISAEDAERWERKKKRKNPDLG
						FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
		ļ				KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
1						KQIEKRDKYSRRRPYNDDADIDYINERNAKF
						NKKAERFYGKYTAEIKQNLERGTAV
935	2285	Α	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
'''			552.			QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
1						SOHSSPAPMYSOTFHILVLG
936	2286	A	8032	1	639	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
750	2200	Α.	0032		039	FRESEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
()						· · · · · · · · · · · · · · · · · · ·
						IWDTAGQERFRTITTAYYRGAMGIMLVYDIT
			'			NEKSFDNIRNWIRNIEEHASADVEKMILGNKC
]						DVNDKRQVSKERGEKLALDYGIKFMETSAK
			[*		ANINVENAFFTLARDIKAKMDKKLEGNSPQG
L		<u> </u>	0000	-303		SNQGVKITPDQQKRSSFFRCVLL
937	2287	A	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
938	2288	Α	8052	675	-1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
						PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
		ĺ				ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
				İ		LREYQTRQDQCIYNTTYLNVQRENGTISRYV
						GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
}						GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
						VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	Α	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
						AEQLKWSAELARLGESIMDGKQGGMDGSKP
, ,				`.		AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
						IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
'						CLGGHLSCVKILLKHGAQVNGVTADWHTPL
·			'		. *	FNACVSGSWDCVNLLLQHGASVQPESDLASP
						IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
					·	
			,		ľ	PLYLACENQQRACVKKLLESGADVNQGKGQ
						DSPLHAVARTASEELACLLMDFGADTQAKN
						AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
	ļ				i	CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
040	2200	 _ 	0050		1002	L
940	2290	Α	8058	2	1203	KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI
لـــــــــــــــــــــــــــــــــــــ						ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI
1		1				ANSVVVWVNIQAKTTGYDTHCYILNLAIADL
					•	WVVLTIPVWVVSLVQHNQWPMGELTCKVTH LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS
						RKKMVRRVVCILVWLLAFCVSLPDTYYLKT
1						VTSASNNETYCRSFYPEHSIKEWLIGMELVSV
						VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS
						RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI
1	·				l	PFTCRLEHALFTALHVTQCLSLVHCCVNPVL YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA
						SRVSETEYSALEQSTK
941	2291	A	8059	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS
						PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV
					·	IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	A	8067	278	1262	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS
						VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT
						MMFWYRQQPGQSLTLIATANQGSEATYESGF
1						VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA GROGTYEOYFGPGTRLTVTEDLKNVFPPEVA
						VFEPSEAEISHTOKATLVCLATGFYPDHVELS
				•		WWVNGKEVHSGVSTDPQPLKEQPALNDSRY
1 1		l				CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE
						NDEWTQDRAKPVTQIVSAEAWGRADCGFTS
						ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG
943	2293	Α	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI
						TASERLRRRPRATARLRAHAAPPEPPLAVFAP
	,					PSDRKELLALPVACDPVIASVMSWVQAASLI
1 1						QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA
						EVILNYGRLRGTLSALLSWCHLHNNNSTLINK
						INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL
1						DSIEDMDLCHVVPAEKKIDEAKDERLCENNA
						EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF
944	2294	Α	8073	1	797	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK
1 1						MAATSGTDEPVSGELVSVAHALSLPAESYGN
						DPDIEMAWAMRAMQHAEVYYKLISSVDPQF
	i					LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD
	ľ				l	CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA
						VYISVQDKEGEKGVNNGGEKRADSGEEENT
						KNGGEKGADSGEEKEEGINREDKTDKGGEK
945	2295	A	8074	2	505	GKEADKEINKSGEKAM GAATLLRSASSAARKAAEAEQVWLHLHRYL
	دوسه	^	3074	-	505	SADRRVLGLREWGRPASERECSLCQRLKREL
						NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT
					ł	GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW
		.				GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER
946	2296	A	8081	42	590	ADLIAYLKKATNE EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF
-			300,			VAIFAVPLILGQEYEDEERLGEDEYYQVVYY
	}					YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK
						DITEAIETTISLETARADHPKPVTVKPVTTEPQ
	}	,			}	SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR
947	2297	A	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF
						SRARAPAHSLRAALSLASSARSWGAVSRDRG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
948	2298	B	8093	3905	846	PCPPAIMYQSSNKC MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHFTRREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEEASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGEGQTPRPLEDTTREAQELEAQLSLVRE GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS GVDEGQGSPSEMVHSSEFRVDNNHLLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKQFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTTTKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY QVDLPHTAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A	8095	9	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKLAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKRHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NLLRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVVMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHVYDLPVEGNEPPASFHKDTNALPPTVFKD NPMSTPKDSNPKA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted begiming nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133	66	1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPRNLLEILLINIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFTYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATTPDNASQEELMITLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS LTMYGALYLYERLSWITHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DRYTTSRRLDPIPOLKCVGGTAGCDSYTPKVI
						QCQNKGWDGYDVQWECKTDLDIAYKFGKT VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL GLQKLKESGKQHGFASFSDYYYKWSSADSC NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ NTGHGATSGFGSAFTGQQGYENSGPGFWTGL GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA SGYGGTRRR
957	2307	A	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA HNEVGTAQCNVTLEVSTOPGAAVVAGAVVG TLVGLGLLAGLVLLYHRRGKALEEPANDIKE DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS QAGSLV
958	2308	A	8161	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP LNATLVITFEITFRSKNITILELPDEVVVPPGVT NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	Α	8163	521	1345	GERAGRRRGRLGVWAQPQPLLPRPVGSRRE MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI ASPDVKLNLGGDFIKESTATTFLRQRGYGWL LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK KPLLIYPIFLLYIYFLSLYTGV
960		A	8167	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN LYSQLNALQFTVDERSILWLNQFLLDLKQSL NQFMAVYKLNDNSKSDEHVDVRVDGLMLK FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI LSSTLTSKGNETIESIFKAEDLLPEAASLSENL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			}	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		1	1	sequence		nucleotide insertion
						DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP
			<u>'</u>			LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL
		}	}			ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN
			}		!	AGANLQNYGETSPDAISTNSEGAQENHDDLM
			•			SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP
]]	DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE
1				·		ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST
						EFLTSSLMNIQHFLEDETVATVMPMKIQVSNT
						KINLKDDSPRSSTVSLEPAPVTVHIDHLVVER
1		}				SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL
					l	TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN
			}			FPEFSFDFTREQLMEENESLKQELAKAKMAL
						AEAHLEKDALLHHIKKMTVE
961	2311	Α	8172	1442	682	TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI
ĺ						ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN
						VSKGQVAKKEDLISAFGTDDQTEICKQILTKG
Į						EVQVSDKERHTQLEQMFRDIATIVADKCVNP
					·	ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA
J						LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL
						KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI
062	2212	<u> </u>	0175	200	507	DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	Α	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS
						VLRRMQKKYWKTKQVFIKATGKKEDEHLVA
						SDAELDAKLEVFHSVQETCTELLKIIEKYQLR
963	2313	_	8181	13	2215	LNGMKS
903	2313	Ą	0161	13	2213	AEGCAERRGTEPVVELSMSWESGAGPGLGSQ GMDLVWSAWYGKCVKGKGSLPLSAHGIVV
ļ						AWLSRAEWDQVTVYLFCDDHKLQRYALNRI
						TVWRSRSGNELPLAVASTADLIRCKLLDVTG
						GLGTDELRLLYGMALVRFVNLISERKTKFAK
	.					VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI
ŀ					_	NDCRRGCYFVLDWLQKTYWCRQLENSLRET
					•	WELEEFREGIEEEDQEEDKNIVVDDITEOKPE
						PODDGKSTESDVKADGDSKGSEEVDSHCKK
1						ALSHKELYERARELLVSYEEEQFTVLEKFRYL
1						PKAIKAWNNPSPRVECVLAELKGVTCENREA
			1			VLDAFLDDGFLVPTFEQLAALQIEYEENVDL
						NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE
					,	RMLSELPALGISGIRPTYILRWTVELIVANTKT
						GRNARRFSAGQWEARRGWRLFNCSASLDWP
1						RMVESCLGSPCWASPQLLRIIFKAMGQGLPD
						EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK
]					SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ
						GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE
	•					EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA
						ENARLLAQKRGALQGSAWQVSSEDVRWDTF
]						PLGRMPGQTEDPAELMLENYDTMYLLDQPV
						LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS
						NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE
						GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR
						PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP
				1		NVGKSTFFNVLTNSQASAENFPFCTIDPNESR
						VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG
1						LVKGAHNGQGLGNAFLSHISACDGIFHLTRA
-[·		.		FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI
			·	į	1	GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS
						WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS
1 1					i	KPMVYLVNLSEKDYIRKKNKWLIKIKEWVD KYDPGALVIPFSGALELKLOELSAEEROKYLE
						STUPENDE VIPENIALE KILIBE NABERLIK VI.E. I

SEQ ID	OFO ID	N.f.	CEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	noa.	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isolaucina, K=Lysina, L=Laucina,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	201100		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	•	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
-				•		ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV
- 1						RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV
1						MKYEDFKEEGSENAVKAAGKYRQQGRNYIV
l					i	EDGDIIFFKFNTPQQPKKK
965	2315	Α	8195	1437	594	RSFSLSFSLLSPSEMMALGAAGATRVFVAMV
ļ						AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL
L						GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY
i						QPYPCAEDEECGTDEYCASPTRGGDAGVQIC
						LACRKRRKRCMRHAMCCPGNYCKNGICVSS
İ					1	DQNHFRGEIEETITESFGNDHSTLDGYSRRTT
						LSSKMYHTKGQEGSVCLRSSDCASGLCCARH
İ						FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ
[RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH
966	2316	A	8207	416	4082	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW
İ						SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI
						DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY
- [WVDLERQLLQRVFLNGSRQERVCNIEKNVSG
1						MAINWINEEVIWSNQQEGIITVTDMKGNNSHI
						LLSALKYPANVAVDPVERFIFWSSEVAGSLY
ļ						RADLDGVGVKALLETSEKITAVSLDVLDKRL
		· ·				FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ
1						HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG
						KDMVRINLHSSFVPLGELKVVHPLAQPKAED
		1				DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC MCAEGYALSRDRKYCEGNDWKYCEDVNEC
ł	·					AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD
1						GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF
Ì						CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL
1						SPVSWECDCFPGYDLQLDEKSCAASGPQPFL
į						LFANSQDIRHMHFDGTDYGTLLSQQMGMVY
[ALDHDPVENKIYFAHTALKWIERANMDGSQ
i			į į			RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK
1						SLIGRSDLNGKRSKIITIENISOPRGIAVHPMAK
1						RLFWTDTGINPRIESSSLOGLGRLVIASSDLIW
						PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK
j						RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP
l						SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP
						LAKPGADPCLYQNGGCEHICKKRLGTAWCS
						CREGFMKASDGKTCLALDGHQLLAGGEVDL
İ					•	KNQVTPLDILSKTRVSEDNITESQHMLVAEIM
.						VSDQDDCAPVGCSMYARCISEGEDATCQCLK
						GFAGDGKLCSDIDECEMGVPVCPPASSKCINT
				,		EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS
ł						CGENASCTNTEGGYTCMCAGRLSEPGLICPD
1		İ				STPPPHLREDDHHYSVRNSDSECPLSHDGYCL
						HDGVCMYTEALDKYACNCVVGYIGERCQYR
						DLKWWELRHAGHGQQQKVIVVAVCVVVLV
	.					MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS
						RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD
						LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ
	[LCGMGTEQGCWIPVSSDKGSCPQVMERSFH
					· !	MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL
067	7217		9210	3	401	DPPHQMELTQ
967	2317	A	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL
						RLHHRFRALDRNKKGYLSRMDLQQIGALAV
						ADDI CODDIDECEDDO COORDO COMO COMO COMO COMO COMO COMO COMO CO
						NPLGDRIESFFPDGSQRVDFPGFVRVLAHFRP
						NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY
						NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL
						NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
968	2318	A	8211	2	409	ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFIT YMDNWRQNTTAEQEALQAKVDAENFYYVIL YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP YHQYIVEDWQEKYKSQILNLEESKATIHENIG AAGFKMSP
969	2319	A	8215	1	1938	GMPRSRGGRAAPGPPPPPPPPPQAPRWSRWR VPGRLLLLLPALCCLPGAARAAAAAAGAGN RAAVAVAVARADEAEAPFAGQNWLKSYGY LLPYDSRASALHSAKALQSAVSTMQQFYGIP VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR NKRYALTGQKWRQKHITYSIHNYTPKVGELD TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL VAVHELGHALGLEHSSDPSAIMAPFYQYMET HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRRIHSPSERKHERQPRPPRPPLGDRPSTPGT KPNICDGNFNTVALFRGEMFVFKDRWFWRL RNNRVQEGYPMQIEQFWKGLPARIDAAYER ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG ELGSCLPREGIDTALRWEPVGKTYFFKGERY WRYSEERRATDPGYPKPITVWKGIPQAPQGA FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY PRNILRDWMGCNQKEVERKERRLPQDDVDI MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN YELLEKEVEKRKVALQEAKLKAKGLNPDGTP ALSTLGGFSPASKPSSPREVKAEEKSPISINVK TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR SASRSRSRTSRSRSHTPRRHYNNRRSRSGTY SSRSRSRSRSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA KKHRHERGHHRDRRERSRSFERSHKSKHHGG SRSGHGRHRR
971	2321	A	8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN VEVEQESFFEGKNMALFEEEMDSNPMVSSLL NKLANYTNLSQGVVEHEEDEESRREAKAPR MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG VLESFLIVAMCCTCTMLTAISMSAIATNGVVP AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA AMLHNMRVYGTCTLVLMALVVFVGVKYVN KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV CLLGNRTLSRRSFDACVKAYGIHNNSATSAL WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA ASGVFLENLWSTYAHAGAFVEKKGVPSVPV AEESRASTLPYVLTDIAASFTLLVGIYFPSVTG IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM LAWPSPWVIVIGSFFSTCGAGLQTLTGAPRLL QAIARDGIVPFLQVFGHGKANGEPTWALLLT VLICETGILIASLDSVAPILSMFFLMCYLFVNL ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL CLALMFICSWYYALSAMLIAGCIYKYIEYRG AEKEWGDGIRGLSLNAARYALLRVEHGPPHT KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

				residue of peptide sequence	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMAWPASWKQEDNPFSW KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD ENYMEFLEVLTEGLNRVLLVRGGGREVTITYS
972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ REKANGTTVHVGIHPSKVVITRLKLDKDRKKI LERKAKSRQVGKEKGKYKEELIEKMOE
973	2323	A	8237	873	4610	GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG AKAWIMDEEEDAEEEGAGGRQDPSRRSIRLR PLPSPSPSAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRRFRGSLASLGSRGGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQOPPQPASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIIHPYSDFRFYWDLTML LLMVGNLIIIPVGITTFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHRYQGKMFDEESILGELSEPLREEINFNC RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE TKLADGSYFGEICLLTRGRRTASVRADTYCR LYSLSVDNFNEVLEEYPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ QIVQHDREMAHCAHRVQAAASATPTPTPVIW TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR PPPGSGLGNLGAGGTPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA GFGHFHKALGGSLSSDSPLLTPLQPGARSPQ AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP PSLVAGASGGASPVGFTTRGGLSPPGHSPGPP RTFPSAPPRASGSHGSLLLPFASSPPPPQVPQR RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGGSGS GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL
974	2324	A	8247	279	468	EYKQWERRFLSCONRNDLGYGKPRKGGGLL LVPVKDASRICSLTYLLGSHWNNLVVRSPVL G

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible nucleotide insertion
975	2325	A	8249	62	1571	LVALKNWKPKGTNIPAPQSPVFGEAVSGVYM MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT ILPQELQAWVQEHCPESAEEAVTLLEDLEREL DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ DPRKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ EAGSKKGRESVPTKPTPGERRYICAECGKAFS NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS NLTLHYRTHLVDRPYDCKCGKAFGQSSDLLK HQRMHTEEAPYQCKDCGKAFSGKGSLIRHYR IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA
976	2326	A	8257	298	7086	GNMACWPQLRLLLWKNLTFRRRQTCQLLLE VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM PSAGTLPWVQGIICNANNPCFRYPTTGEAPGV VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS GFLYHNLSLPKSTVDKMLRADVILHKVFLQG YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA VFHDLEGMWEELSPKIWTTMENSQEMDLVR MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER KFWAGIVFTGITTGSIELPHHVKYKIRMGIDN VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV FLSVFAVVTILQCFLISTLFSRANLAAACGGII YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE EDGFNLTTSVSMMLFDTFFLYGVMTWYIEAVF PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD GMKVAVDGLALNFYEGQITSFLGHNGAGKT TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ NLGVCPQHNVLFDMILTVEEHIWFYARLKGLS EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP YSRRGIWELLLKYRQGRTIILSTHHMDEADVL GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT LVKKDVESSLSSCRNSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID DRLSDLGISSYGISETTLEEIFLKVAEESGVDA ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA ADPNDSDIDPESRETDLLSGMDGKGSYQVKG WKLTQQQFVALLWKRILLIARSRKGFFAQIV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	l	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ŀ	peptide	ļ	/=possible nucleotide deletion, \-possible
	<u> </u>			sequence		nucleotide insertion
						LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY
					j	NEQYTFVSNDAPEDTGTLELLNALTKDPGFG
						TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM
ŀ						DLFQNGNWTMQNPSPACQCSSDKIKKMLPV
J		J	,		J	CPPGAGGLPPPQRKQNTADILQDLTGRNISDY
			ļ			LVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVS
						NTQALPPSQEVNDATKQMKKHLKLAKDSSA
		l				DRFLNSLGRFMTGLDTRNNVKVWFNNKGW
						HAISSFLNVINNAILRANLQKGENPSHYGITAF NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA
	1	1			[MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI
1						YWLSNFVWDMCNYVVPATLVIIIFICFOOKSY
		1	1			VSSTNLPVLALLLLLYGWSITPLMYPASFVFK
j]				PSTAYVVLTSVNLFIGINGSVATFVLELFTDN
						KLNNINDILKSVFLIFPHFCLGRGLIDMVKNO
	1	ļ				AMADALERFGENRFVSPLSWDLVGRNLFAM
		1				AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN
			İ			DEDEDVRRERQRILDGGGQNDILEIKELTKIY
					İ	RRKRKPAVDRICVGIPPGECFGLLGVNGAGK
						SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV
		ŀ				HQNMGYCPQFDAITELLTGREHVEFFALLRG
						VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY
ł		ľ				SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD
			•			PKARRFLWNCALSVVKEGRSVVLTSHSMEEC
						EALCTRMAIMVNGRFRCLGSVQHLKNRFGD
	·					GYTTVVRIAGSNPDLKPVQDFFGLAFPGSVPK
1						EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH
					•	IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL
977	2327	Α	8260	3	1567	SLHKNQTVVDVAVLTSFLQDEKVKESYV IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG
]""	2321	^	8200	3	1507	YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV
]						TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM
					•	SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH
1.						YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS
						LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM
						VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV
						MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG
1			.			MYAYAGWFYLNFVTEEVENPEKTIPLAICISM
						AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT
						FSERLLGNPSLAVPIFVALSCFGSMNGGVFAV
			ľ			SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV
						LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA
						VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC
] i						LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII
978	2220		D261	3	2165	WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL
7/0	2328	A	8261	2	2165	RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL
						RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP
						LADAASMSGVRAVRISIESACEKQVHEVGLD
}				l		GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK
				l	ļ	FOPSLWPWDSVRNNLRSALTEMCVLYDVLSI
	· .			-		VRDKKFMTLDPVSQDALPPKQNPQTLQLISK
				İ		KKSLAGAAQILLKGAERLTKSVTENQENKLO
]			}	J		RDFNSELLRLRQHWKLRKVGDKILGDLSYRS
	. 1				ŀ	AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL
						DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN
[·		ļ			LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI
]			ł	İ		FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ
	[[f		[LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE
			·			HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPOIQAHWSNIND VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI
						RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASAITVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ WNKMEGRNFVYKMELLMSALSPCLL
979	2329	A	8289	2	1053	FVWNPRGGRKRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLPVLLMTLFFSGLILAILLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	A	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGGSPGACPAL GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGFRARSNMHFTSSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRRTNTSSVTTTITQSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN
982	2332	A	8315	1	1004	GSTHASTORINGE VIBINITED VALENCIANS KGNEPPPPPI GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEEQKEEEERKAREEQAQREHEEYLKLKEA FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFTYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA
983	2333	A	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF DLGPFIVGSLGPPDLITPTEGSGRSPRYALWFC VGESWPQDQPWTKRLVMVKVVPTCLRALVE MARVGGASSLENTVDLHISNSHPLSLTSDQY
984	2334	A	8321	1	1243	KAYLQDLVEGMDFQGPGBS ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATRRILAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89	1172	KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	A	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RTRIKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	A	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL KSLHPMS
990	2340	A	8361	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

NO: of nucl- cotide seq- uence NO: of peptide seq- uence NO: of poptide seq- uence NO: of peptide seq- uence No: of peptide seq- uence No: of peptide seq- uence No: of peptide seq- uence No: of peptide seq- uence No: of peptide seq- uence No: of peptide seq- uence No: of peptide seq- vence seq-	H=Histidine, ucine,
sequence uence 09/496 correspondi to last amino M=Methionine, N=Asparagine of peptide T=Threonine, V=Valine, W=T	
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S= amino acid of peptide T=Threonine, V=Valine, W=I	
amino acid of peptide T=Threonine, V=Valine, W=T	
peptide /=possible nucleotide deletion,	, \= possible
sequence nucleotide insertion	
LIKKYDQMAIFHCLFWPSL RLTQHLSLLCEKYSTVVRI	ILLUGALIYUMV
QFKLCIMRRSKGRAEKS	EVOGEVETIEQH
991 2341 A 8369 9 921 SSVVEFSALSVSMACLSPS(OI OKEOODCEI VII
EGFLSAEECVAMQQRIGET	
EFSTQEEEQLRAQGSTDYF	
GVFDEKGNFLVPPEKSINK	,
SITHSFKVQTLARSLGLQM	
HFGGEVSPHQDASFLYTEP	
DATLENGCLWFIPGSHTSG	VSRRMVRAPVGS
APGTSFLGSEPARDNSLFVI	PTPVQRGALVLIH
GEVVHKSKQNLSDRSRQA	
WSPENWLQPTAELPFPQLY	
992 2342 A 8370 906 4 MALSGNCSRYYPREQGSA	
GGQVYFTRHSTLISIPHSLL	
DLAKDSKGRFFIDRDGFLF	
LPDHFPEKGRLKREAEYFQ	
KQSPDEFCHSDFEDASQGS RKWGFITVGYRGSCTLGRE	
ILVCGRISLAKEVFGETLNE	•
RFYLKFKHLMGAPASNFIL	
HPVNIYLQQRSVIRPDLTSK	, ,
QEVSRRRRWLGDPEHL	ad to batton ob it
993 2343 A 8379 1 2794 MRMQRHKNDTMDFGDSG	KRIGGGVLCLLHO
SNTSFIKLNNNGFEDIVIVII	
DMVTTASTYLFEATEKRFF	TKNVSILIPENWK
ENPQYKRPKHENHKHADV	IVAPPTLPGRDEP
YTKQFTECGEKGEYIHFTPI	
PPGKLFVHEWAHLRWGVF	•
KSKKIEATRCSAGISGRNRV	
CRIDSTTKLYGKDCQFFPDI	
QSIDSVVEFCNEKTHNQEA	
WEVISNSEDFKNTIPMVTPI	
VCLVLDKSGSMGGKDRLN TVENGSWVGMVHFDSTAT	
NTLMAGLPTYPLGGTSICS(
SQLDGSEVLLLTDGEDNTA	
VHFIALGRAADEAVIEMSK	
QNNGLIDAFGALTSGNTDL	
LNSNAWMNDTVIIDSTVGK	
SISLWDPSGTIMENFTVDAT	
KVGTWAYNLQAKANPETL	TITVTSRAANSSV
PPITVNAKMNKDVNSFPSPI	
VLGANVTAFIESQNGHTEV	
FKNDGVYSRYFTAYTENGF	
NTARLKLRPPLNRAAYIPG	
RPEIDEDTQTTLEDFSRTAS	
PLPDQYPPSQITDLDATVHE NFDVGKVQRYIIRISASILDI	
TTDLSPKEANSKESFAFKPE	
KSIDKSNLTSKVSNIAQVTL	
PTPTPTPDKSHNSGVNIST	
NFILSTTI	75 - 700 1 100 4 41 4
994 2344 A 8385 231 644 INSSPRTGRDHQELNLHTER	DSRSORAVIED
RQNPGIFYWIFLPSRSHSASI	
TODEILKMRNTFAELKNSLI	
RIGTQAGVQWRDHGSLQPQ	
PSSWDYRACLS	` ` ` ` ` `
995 2345 A 8390 194 3421 AWRKSSVVPPRGTRRGEKS	DQDKSGQKNKR

SEQ ID Not of nucleotide per principate Decided of per principate Decided of per principate USSN USSN Usence USSN USS							
peptide coide sequence USSN (99/96 corresponding plat in moteroid plat in mice and preside peptide sequence) ### USSN (19496 corresponding plat in mice and preside peptide sequence) ### USSN (19496 corresponding plat in mice and preside peptide sequence) #### USSN (19496 corresponding plat in mice and preside peptide sequence) #### USSN (19496 corresponding plat in mice and preside peptide sequence) #### USSN (19496 corresponding plat in mice and preside peptide sequence) #### USSN (19496 corresponding plat in mice and preside peptide sequence) #### USSN (19496 corresponding plat in mice and preside peptide sequence) #### USSN (19496 corresponding plati			1				
uence USSN play66 per union programment p			поа	I .		(1
uence 09/496 014 015			1	1			,
uenice	1		į	1			
minto acid residue of peptide reduced peptide sequence pi		ucia			•		
residue of peptide peptide peptide sequence Poptible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelout, hyposible nucleotide detelouted nucleotide detelouted nucleotide nucleoti	1 2000			* * * * * * * * * * * * * * * * * * *			
peptide sequence Poposible nucleotide deletion, li-possible nucleotide insertion poposible nucleotide insertion poposible poposib	1						
Incledité insertion						344233	
DPILSMKQSPALAPEERCRRAGSPKPVLRADD NNMGNGCSQKLATANLRFLLLVIPICICALV LILEILISYVGTLQKVYPKSNGSEPLVTDGEI QGSDVLITNTTYNGSTVSTAHPDQHYPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYH ATLTPILSVVRNMEMERCKSPCQMLPYH ATLTPILSVVRNMEMERCKSPCRMITHSCC AKEGCESVLGMNYSWPDFLRCSOFRNQT ESSNVSRICFSPQENGKQLLGGRGENFLCAS GICIPOKLQCNGYNDCDDWSDEARLGNCSSRIL FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC CDRYCHENCEGRICAMEWVCDGDDCDC DEDCKDGSDEERCSVQTSCOEGOPDCVD KSDEVNCSCHSQGLYECRIGQCPSTTPCCDG DEDCKDGSDEERCSVQTSCOEGOPGCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPTTLELCM MLPYNSTSYPHYFORLTOKEASINSWSSLEPA LVQTNCVKYLMFFSCTTLVPRCDVNTGEHIPP CRALCEHISKERCSSVLGVINGLQWPEDTDCSO FEENDADTCLMPDEVYECCSSNLFKCRSGO CVLASRRCDGQADCDDSDEERCCGKERDL WECPSNKQCLKHTVLCDGFPDCPDYMDEKN SECONDOLLECANIHACVSBLWCDGADCS CSCQDDELECANIHACVSBLWCDGADCS SSSDEWDCVTLSINVRSSSTLAVHRAATEHH VCLOEWGELLSCLACKQGGLGEPSYTLLIGE GEKEPRWITLHSNWSSLAGTHELLVNOQS CEERSKELLLXKQDCGREPARAWKRILGGR TSRRGWPWQCSLSSEPSGHICCCVLLAKW VLTVAHCFEGRENAAWKRVLGGR TSRRGWPWQCSLSSEPSGHICCCVLLAKW VLTVAHCFEGRENAAWKRVLGDNLDHS VMGTRYVCTILLHRSTSRAVDYDSIVLISE DISTOTYVRVCLLPPEQWLEPDTYCTTTGW GHMGNKAPFFLQSECVSTSRAGFYSTVDK GHMGNKAPFFLQSECVSTSRAGFYSTVDK GHMGNKAPFFLQSECVSTSRAGFYSTVDK GHMGNKAPFFLQSECVSTSRAGFYSTVDK SRRGSSSSRSSRSSRSSRSSRSSRSSRSSRSSRSSRSSRS	l	i	1		,		1 •
LILLEILSYVGTLQKYYFKSNGSEPLVTDGEI QGSDVILTNITIYNGSTVSTAHPDQHYPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYH ATLTPILSVVRNMEREKFLKFFTLHRISCY QFIIMLFGCTLAFPCUIDGDDSHGLLPCRSPCE AAKEGCESVLGMVNYSWPDFLACSOFRNQT ESSNVSRICTSPQQENGKQLLCGRCENFLCAS GICIPOKLQCNGYTDCDDVSDBAHCNCSBNL FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD SISBEVNSCSHQGLVSCRNQCQCPSTPCCDG DEDCKDGSDEENCSVQTSCQEDGRCLYNFP CLBSCGGSSLCDPNNSCHNCSPTPLECM NLPYNSTSYPRYFGHRTQKEASISWESSLFPA LVQTNCVKYLMFFSCTLIVKPCLVNYTGEHPP CRALCEHSKERCESVLGNVGLQWPEDTDCSQ PPENSDNQTCLMPDLYVEGCSSIFTCRGG FPENSDNQTCLMPDLYVEGCSSIFTCRGGO CVLASRRCDGQADCDDDSDEENCGCKERDL WCCPSNKQCLKHTVLOGFPDCPDYMDEKN CSFQQDDELECANHACVSRDLWCDGEADCS SDSDEWDCVTLSINVSSSSLWPTAATEHH VCADGWQELLSOLACKQMGLGEPSVTLLIQE QKEPFRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPFRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRWTLTHSNWSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRWTLTHSNWSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRKTLTHSNWSSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRKTLTHSNWSSSLWFTAATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRKTLTHSNWSSSSLWFTAATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE VFWMKRQPTIQTTSGGTSQDTK VFWMKRQPTIQTTSGGTSQDTK VFWMKRQPTIQTTSGGTSQDTK VFWMKRQPTIQTTSGGTSQDTK ASSSEPPBAGNGGNASASAVSELSPER SPALKSPLQSVVVRRSSPSPSRSSSSSSSSSSSSSSSSSSSSSSSSSS					-		DFLSMKQSPALAPEERCRRAGSPKPVLRADD
LILLEILSYVGTLQKYYFKSNGSEPLVTDGEI QGSDVILTNITIYNGSTVSTAHPDQHYPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYH ATLTPILSVVRNMEREKFLKFFTLHRISCY QFIIMLFGCTLAFPCUIDGDDSHGLLPCRSPCE AAKEGCESVLGMVNYSWPDFLACSOFRNQT ESSNVSRICTSPQQENGKQLLCGRCENFLCAS GICIPOKLQCNGYTDCDDVSDBAHCNCSBNL FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD SISBEVNSCSHQGLVSCRNQCQCPSTPCCDG DEDCKDGSDEENCSVQTSCQEDGRCLYNFP CLBSCGGSSLCDPNNSCHNCSPTPLECM NLPYNSTSYPRYFGHRTQKEASISWESSLFPA LVQTNCVKYLMFFSCTLIVKPCLVNYTGEHPP CRALCEHSKERCESVLGNVGLQWPEDTDCSQ PPENSDNQTCLMPDLYVEGCSSIFTCRGG FPENSDNQTCLMPDLYVEGCSSIFTCRGGO CVLASRRCDGQADCDDDSDEENCGCKERDL WCCPSNKQCLKHTVLOGFPDCPDYMDEKN CSFQQDDELECANHACVSRDLWCDGEADCS SDSDEWDCVTLSINVSSSSLWPTAATEHH VCADGWQELLSOLACKQMGLGEPSVTLLIQE QKEPFRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPFRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMAATEHH VCADGWQELSOLACKQMGLGEPSVTLLIQE QKEPRWTLTHSNWSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRWTLTHSNWSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRWTLTHSNWSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRKTLTHSNWSSSSLWFMATATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRKTLTHSNWSSSLWFTAATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE QKEPRKTLTHSNWSSSSLWFTAATEHH VCADGWQELSOLACKQMGLGPSVTLLIQE VFWMKRQPTIQTTSGGTSQDTK VFWMKRQPTIQTTSGGTSQDTK VFWMKRQPTIQTTSGGTSQDTK VFWMKRQPTIQTTSGGTSQDTK ASSSEPPBAGNGGNASASAVSELSPER SPALKSPLQSVVVRRSSPSPSRSSSSSSSSSSSSSSSSSSSSSSSSSS							NNMGNGCSQKLATANLLRFLLLVLIPCICALV
TDASLPEGOSHRITISACAMITHSQCQM_PYI AILTPLLSVVRNMEMEKILKPTYJHRLSCY QHIMLFGCTLAPPECIIDDDSHGLLPCRSEC AAKEGCESVLGMVYSWPDPLRCSQFRNQT ESSNVSRICTSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGVHYDCDDWSDEATHCASEN HECHTGKCLNYSLVCDGVDDCCDLSDEDCVD ESSNVSRICTSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGVHDCDCDLSDEDCVD LSDEWNCSCHSQQLVECNNQCGPSTTQCDB DEDCKDGSDEENGSVIGTSQEGDQRCLYMP CLDSCGGSSLCDPNSLNNCSQCEPTIELCM NLFYNSTSYPNYFGRTQKEASISWESSLFT LQTITACYXYLMFFSCTILLYRCCHOVTGEHIPP CLDSCGGSSLCDPNSLNNCSQCEPTIELCM NLFYNSTSYPNYFGRTQKEASISWESSLFT CRALCEHSKERCESVLGIVGLQWPEDITDCSQ FPEENSNQTLCHPDEVYWECSPSHFKCRSQQ CVLASRRCDQADCDDDSDEENCGCKERDL WECPSNKQCLKHVICOGFPDCTYMDERN CSFQQDELECANHACVSRDLWCDGEADCS FPEENSNQTLCHVICOGFPDCTYMDERN CSFQQDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSIFMHAATEHH VCADGWCGLSQLACKQMGLGEPSVTLLIQE QEKEPRWTLTLSNNVSSSIFMHAATEHH VCADGWCGLSQLACKQMGLGEPSVTLLIQE QEKERRWTHINFNSSNSTNSTSSTSSTSSTSSTSSTSVSPAHN CSFQQDELECANHACVSRDLWCDGAACC GESRSKISLLCTKQDCGRRAARMKKRLGCN TSRCGRWPWQCSLSQSEPSGHICGCVLLAKKW VLTVAHCFGGRENAAVWKVLGINNLDHPS VPMQTFFKYTLHTSPSSAVVVVLGINNLDHPS VPMQTFFKYTLHTSPSSAVVVVLGINNLDHPS VPMQTFFKYTLTPSQWLEPDTYCYTTOW GHMGNMAPFKLQGGFVSILSHELCOSYFDMK TITTRMICAGYESGTVDSCMGDSGGFLVCEK PGGRWTTGLTSWGSVCPSKVLGPGVYSTMW GHMGNMAPFKLQGGFVSILSHCOSYFDMK TITTRMICAGYESGTVDSCMGDSGGFLVCEK PGGRWTTGLTSWGSVCPSKVLGPGVYSTMWNNYNQAY SPRRGRSRSSRSSKSKSKSSSSSRSRSSSSSSSSSSSSSS			ļ [.]	}			
ATLTPLISVVRNMEMEKFLKFFTYLHRILSCY QHMLIFGGTLAFPGEIDGDDSHGLLQCRSPGL AAKEGCESVLGMYNYSWPDFLRCSQFRNOL ESNNSRIGFSPGOENGKOLLGGGENFILCAS ESNSRIGFSPGOENGKOLLGGGENFILCAS ESNSRIGFSPGOENGKOLLGGGENFILCAS ESNSRIGFSPGOENGKOLLGGGENFILCAS ESNSRIGFSPGOENGKOLDGGENFILCAS ESNSRIGFSPGOENGKOLDGDHDCVD KSDEVNSCSHSQGLVECRMGCUFSTFQCOB DEDCKGSDEENGSVIGTSCGGDORGCLYDPI CLDSCGGSSLCPRNSLNNCSQCEPTITELCH MLPYNSTSYPNYTGHRTQKASISWESSLFPA LVQTNCYXYLMFFSCTILVPKCDVNTGEHPP CRALCEHISKERGESVLGIVGLQWPEDTDCSQ PFERNSDNQTCLMPDEYVEGCSRSHFKCRSQ QCVLASRRCDGGADCDDDSDEENGCCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDELECAMHACVSRDLWCGGEADCS DSSDEWDCVTLSINNNSSSTLMYHRAATEHH VCADGWGELSJCLACKQMGLGFPCCVTLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRKSLLTKQDCGTRAARMKKILGGG RTSRFGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAWKVVLGINNLDHSP VTMQTFFVKTILHFRYSKAVVDYDISIVELSE DISSTGYVRPVCLINPEQWLEPDTYCYTTGG GHMGNMPFKLGGGEVSILSLELCQSYFDMK TITTRMCAGYESGTVDSCMGDSGGPLVCK PGGRWTLFGLTSWGSVCFSWLOFGCYSNSVS YFVEWIKRQUTJOFFLLN RICHARDS SSRSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	1		l	1			QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT
QHIMLFGCTLAPPECIIDODSHGLIPCRSEQ AAKEGCESVLGWNYSWPDFIRESOPRINQT ESSNYSRICTSPQOENGKOLLGRGENFILCAS GICIPGKLQCNGYNDCDDWSDEAHCNSENL FICHTGKCLAYSLVCDGYDDCGDLSDEQNC DCNPTTEHRGGDGRCLAMEWCDGDHDCVS KSDEVNOSCHSQGLVECNRQCCPSTTQCDG DEDCKDGSBEENCSVIGTSCQEGDQRCLYNP CLDSCGGSSLCDPNSLNNCSQCEGTTLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGWGLQWPEDTDCSC FIENSNONTCLMPBCYTECSSPSIFECRSQQ CVLASRRCDGADCDDDSDEENCGCKERDL WECPSIKQCLKHTVICDGFPDCPTYMDEKN CSFQQDDELECANHACVSRDLWCDGBADCS DSSDEWDCVTLSINNYSSSELWYHRAATEHH VCADGWQELISQLACKQMGLGEPSVTKLIGG QEKEPRWLTLHISNWESLANGTILHELLVNQGS CESSISISLLCTKQDGCRRPAARMKKRLGGR TSRFGRWPWQCSLQSEPSGHIGGCVLLAKKW VLTVAHCPEGGERAAAWKKVLLGGR TSRFGRWPWQCSLQSEPSGHIGGCVLLAKKW VLTVAHCPEGGERAAAWKKVLLGGR TSRFGRWPWQCSLQSEPSGHIGGCVLLAKKW VLTVAHCPEGGERAAAWKKVLLGGR TSRFGRWPWQCSLQSEPSGHIGGCVLLAKKW VLTVAHCPEGGERAAAWKKVLLGGR TSRFGRWPWQCSLQSEPSGHIGGCVLLAKKW VLTVAHCPEGGERAAAWKKVLLGGRPLAKKW TITTRIMCAGYESGTVDSCMGDSGPLVCK PGGRWTLFGLTSWGSVCPSKVLGPGVYSNVS YFVEWIRKQNTJOFFQCMEDPTCYTTGW GHMGNKMPFKLQGGEVEISLEHCGSYPDNK TITTRIMCAGYESGTVDSCMGDSGPLVCK PGGRWTLFGLTSWGSVCPSKVLGPGVYSNVS YFVEWIRKQNTJOFFGLHRGYRYFYFRGR NRGFYPWQOYNRGGYGRYRSNWQNYXQAY SPRRGSTSSFSKRRSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS		1					TDASLPGDQSHRNTSACMNITHSQCQMLPYH
AAKEGCESVLGMYNYSWPDFLRCSQFRNQT ESSNYSRICPSPQORMGQLOGREPHCAS GICIPGKLQCNGYNDCDDWSDRAHCNCSENL FHCHTGKCLINYSLVCDGYDGLSDEQNC DCMPTTEHRCGDGRCIAMEWVCDGDHDCVD KSDEVNCSCHSQGLVECRNQCJESTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNSLNNCSQCEPTITELCM NLPYNSTSYPHYGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHPP CRALCEHSKERCESVLQIVGLQWPEDTDCSQ PFEENSDNQTCLMPDEVVEECSPSHFKCRSGQ CVLASRRCDGQADCDDDSDENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSSCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLWVHRAATEHH VCADGWGELLSQLACKQMGLGPSTVEKLOG QEKEPRWLTLHSNWESINGTTLHELLVNOQS CESRSKISLLCTKQDCGRPAMKKRLIGGR TSRPGWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGNNLIDHS VFMQTRFVKTILHPRYSRAVVDVDISIVELSE DISETGYVRPVCLPNPEQWLEEDTYCYTTGW GHMONKMPFKLOGEVFUNSLEHCQSVFDMK TITTRMCAGYBSGTVDSCMGDSGGPLVCEK PGGRWTLFGLISWGSVCPSKVLOPGVYSNVS VFVEWKRQTVQTFLLN 996 2346 A 8392 199 3085 KVLSSEMSTINTSKSGSGSSSRSSASSRSRSSS SRSSSSSSSSSSSSSSSSSS	1	1	[
ESSNYSRICFSPQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDUSDABAHCNCSENIL, FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHECGGRCIAMEWVCDGDHDCVD KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPTITELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYXVLMFFSCTLLVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEVVEECSPSHFKCRSQ CVLASRRCDGQADCDDBDEENCGCKERDL WECPSNKQCLKHTVICDGPTOCYMDEXN CSPCQDDBLECANHACVSRDLWCDGBADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWGELSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTHELLVNQGS CESRSKISLLCTKQDCGRPAARMNKRLIGGR TSRPGRWPWQCSLQSEPSCHICGCVLIAKKW VLTVAHCFEGENAAVWXVIDISIVISLSE DISETGYVRPVCLYPHPGWLEPDTYCVTTGW GHMGNKMPFKLQSGEVRISLEHCQSYPDMK TITTMCAGYPEGNAAVWXVIDISIVISLSE DISETGYVRPVCLYPHPGWLEPDTYCVTTGW GHMGNKMPFKLQSGEVRISLEHCQSYPDMK TITTMCAGYPEGNAAVWXVIDISIVISLSE DISETGYVRPVCLYPHPGWLEPDTYCYTTGW GHMGNKMPFKLQSGGVRSKVLGPGVYSNVS YFVEWIKRQTVIQTFLLN FSKSSKSRSLSSRSKRSLSSRSRSSRSSSSSPSSSSSSSSSS							l =
GICIPGKLQCNGYNDCDDWSDEAHCNCSENL FHICHTIGKCLINYSLVCOGYDGLSDEQNC DCNPTTEHRCGDGRCIAMEWVCIODHDCVD KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNF CLDSCGGSSLCDPNNSLNNCSQCEPTITELCM NILPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ CPVLASRRCDGADCDDDSBECGCKERDL WECPSNKQCLKHTIVICOGPPDCPDVMDEEN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSLTMVHRAATEHH VCADGWQELLSQLACKQMGLGEPSVTKLIGE QEKEPRWLTLISINWESLINGTLHELLVNQQS CESSKSISLLCTKQDCGRRPAARMNKRLIGGR TSRFGRWPWQCSLQSEPSGHTCHLLVNQQS CESSKSISLLCTKQDCGRRPAARMNKRLIGGR TSRFGRWPWQCSLQSEPSGHCVLIAKKW VLTVAHCFEGRENAAVWXVVLGINILDHPS VFMQTRFVKTILHPRYSRAVDVISIVELSE DISETGYVRPVCLPNFEQWLEPDTYCYTTGW GHMGNKMPFKLQGEVRIJSLEHCQSYFDMK TITTRMCAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YVFWWKRQTVJQTTLH SKSSKSSRSSSKSSRSSRSSRSSRSSRSSSSSSSSSSS	1		1				l
FHCHTOKCLNYSLVCDGYDDCGLDSDEONC DCMPTTEHRCGGRCIAMEWCDGDHDCVD RSDEVNCSCHSQGLVERNGQCPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFA LVQTINCYXLMFFSCTILVPKCDVNTGEHIPP CRALCHSKERCESVLGIVGLQWFEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSQ CVLASRRCDGQADCDDDBDEENCGCKERDL WECPSNKQCLKHTVICDGFPCDYMEISIN CSFCQDDELECANHACVSRDLWCDGEADCS DSDEWDCVTI,SINVNSSSTHWRATEHH VCADGWQEILSQLACKQMGLGEPSYTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRAARAMKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGENAAVWXVDYDISIVELSE DISETGYVRPVCLPPFEQWLGPDTYCYTIGW GHMGNKMPFKLQSGEVRISLEHCQSYFDMK TITTMICAGYESGTVDSCMGDSGQPLVCEK PGGRWTLFGLTSWGSVCPSKVLQPGVYSNVS YFVEWIKRQVIOTHLIN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSSRSASRSRSRS FSKSRSRSSLSSRKRKLSSRSRSSYSPAIN RENHPRVYQNRDFRGHNRGVRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWNYNQAY SPRRGRSSRSPKRSSPSFRRSPSPSSPAIN RENHPRVYQNRDFRGHNRGVRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWNYNQAY SPRRGRSSRSPKRSSPSFRRSPSPSPSPLSS SAKSPVGAAGNQGDEVECGTFGGGTGQDTK ASESSKPWPDATYGTGSASRASAVSELSPRE SPALKSPLGSVVRRSSPFSPSPLSSPRE SPALKSPLGSVVRRSSPFSPSPLSSPRE SPALKSPLGSVVRRSSPFSPSPLSSPRE SPALKSPLGSVVRRSSPFSPSPSPLSSPRE SPALKSPLGSVVRRSSPFSPSPLSSPRE SPALKSPLGSVVRRSSPFFGFTGGGGGDHGGGFTGGGGTGGDTK ASESSKPWPDATYGTGSASRASAVSELSRFRG SSSRSPSGAGGGGFTGGGGGTGGDTK ASESSKPWPDATYGTGSASRASAVSELSFRE SPALKSPLGSVVRRSSPFFGFTGGGGGGGGGGGFTGGGGGTGGDTK ASESSKPWPDATYGTGSASRASAVSELSERE SPALKSPLGSVVRRSSPFFGFTGSGGGGGGGGFTGGGGGGGGGGGGGGFTGGGGGGGG	1		ļ				l
DCMPTTEIRCGDGRCIAMEWVCDGDHDCVD KSDEVNCSCHSQGI-VERNOQCPSTTPQCDG DEDCKDGSDEBNCSVIQTSCQEGDGRCI_YNP CLDSCGGSSLCDPNNSI.NNCSQCEPTITELCM NLFYNSTSYPNYFOHRTQKEASISWESSLFA LVQTNCYKYLMFFSCTILVPKCDVNTGEHPP CRALCEHSKERCESVLGIVGIQWPEDTDCSQ FPEENSDNQTCLMPDEVVEECSPSHFKCRSQ CVLASRRCDGADCDD5BDEBNCGCKERDL WECPSNKQCLKHTVICDGFPDCPDVMDEKN CSRCQDDELECANHACVSRDI_WCDGBADCS DSDEWDCVTI.SINVNSSSFLMVHRAATEHH VCADGWGELSQLACKQMGI.GEPSVTKLIGE QEKERWLITLHSNWESLINGTI.HELLVNKQOS CESRSKISLLCTKQDCGRRPAARMNKRILGGR TSRPGNEPWQCSLQSEPSGHTCHLLVNKQOS CESRSKISLLCTKQDCGRRPAARMNKRILGGR TSRPGNEPWQCSLQSEPSGHCCVLIAKKW VLTVAHCFEGRENAAVWKVVI.GINI.DHPS VFMQTFRV*CII.HPRYSRAVDVISIVELSE DISETGYVRPVCI.PNFEQWLEPDTYCVTITGW GHMGNKMPFKLQGEVRIISLEHCQSYFDMK THTTMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTI-FGLTSWGSVCFSKVI.GPGVYSNVS YFVEWIKAQTVQTTLLN THTTMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTI-FGLTSWGSVCFSKVI.GPGVYSNVS YFVEWIKAQTVQTTLLN THTTMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTI-FGLTSWGSVCFSKVI.GPGVYSNVS YFVEWIKAQTVQTTLLN THTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTI-FGLTSWGSVCFSKVI.GPGVYSNVS YFVEWIKAQTVQTTLLN THTTMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTI-FGLTSWGSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	1	ł	1	ł			
KSDEVNCSCHSQQLVECRIGQCPSTIQCDG							
DEDCKDGSDEENCSVIJOTSCGEORCILYNF CLDSCGGSSLCPNNSINGSCPFITELLCM NLPYNSTSYPNYFGHTQKEASISWESSLFPA LVQTINCYKYLMFFSCTILVPKCDVITGEHIPP CRALCEHISKERCESVLGPIGUGUWPEDTDCSQ FPEENSDNOTCLMPDEYVECCSPSHFKCRSGQ CVLASRRCOGADCDDDSENOGCKERDL WECPSNRQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINNVISSSTLMYHRAATEHH VCADGWGELSQLACKQMGLGEPSVTKLIQE QEKEPRWILHISMWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMYKRILGGR TSRPGRWPWQCSLOSEPSICGCVLLAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRTVKTITLHPRYSRAVVDYDISVELSE DISETGYVRPVCINPROGUEOPDTYCYTTGW GHMGNKMPFKLQGEVRIISLEHCQSYFDMK TITTRMICAG YESGTVDSCHGCYVLIAKKE GHMGNKMPFKLQGEVRIISLEHCQSYFDMK TITTRMICAG YESGTVDSCHGCYVLIAKKE STENGRWPYQCINPROGSGPLVCKE PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLIN **VFWGRINGTHIND*** **VFWGRNTHFRYSTRAVTYPTYFRGR NRGFYPWQTYNRGGYGNYRSNWONYRQAY SPRRGRSRSSRSSRSSSSSSSSSSSSSSSSSSSSSSSSS	1						
CLDSCGGSSLCDPNNSINNCSQCEPITLELCM NLPYNSTSYPNYFQRFROKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ PPEENSDNTCLMPDEYVPECSPSHFKCRSGQ CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDEPCDPYMDEKN CSFCQDDELECANHACVSRDL WCDGBADCS DSSDEWDCVTLSNNVSSSTLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWITLHSNWESLNGTTHELVNGQS CESRKISLLCTKQDCGRRPAARMNKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFGGRENAAVWVJGINNLDHPS VFMQTRFVKTITLHPRYSRAVVDYDISIVELSE DISETGYVPVCLPNPBQWFTPCYTTGW GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPOVYSNVS YFVEWIKRQTYJOTFLLN KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS PSKSRSSRSLS,SRSRKRALSSRSRSRSSPSPAHN RERNHPRVYQNRDFRGHNRGYRPYYFRGR NRGFYPWQQYNRGGYGNYRSNWQNYRQAY SPRGRSSRSSNSHSRYSSSSNSNSNDKSSS DRSRSSSSSSSSSSNSNSNDKSSS DRSRSSSSSSSSSSNSNSNDKSSS DRSRSSSSSSSSSSNSNSNDKSSS NDSRSSSSSSSSSSSNSNSNDKSSS NDSRSPQAAGDNQGDEVKEQTTSGGTSQDTK ASSSKPWPDATYGTGSARASAVSELSPRE SPALKSPLQSVVVRRRSPRSPSPVPKPSPLSST SQMGSTLPSGAGYQSGTHQQGPDHOSGSLSP SKLSPYQSSYPGAGYQSGTHQGGPDHOSGSLSP SKLSPYQSSPPGAGYQSGTHQGGPDHOSGSLSP SKLSPYQSSPPGKSSFPGKSTSSRSPRSSPPRFRG SQSPKRYKLRDDFEKKMADFHKEEDDQDK DKAAGGRKESFDDPEFKFMSVIGANKNQEE KSGKWGLVYAPPGKEKQAFTNTSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEDDDQDK DKAAGGRKESFDDPEFKFMSVIGANKNQEE KSGKWGLVYAPPGKEKQAFTNTSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEDDDQDK SSSLEFGDGKKMSSDFAFTTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEDDDQDK SSSLEFGDGKKMSSCGHRGFVPEKNRTVTAYK AVQEKSSSPPFRKTSSENDKLGAKGDFFTGKS SSSITRBAQVNVRNDSFDEDLARPSGLAQEK KLCRDLVMSNKEQGEFRSDHQSAQSQSPS SELFAQHVTVTVHHVKEHHFGSSGMLHERFT KYLKRGTGQEAAKNKKSPEHRRDISSFSTFRK HGLAHDEMKSPREPGYKARGKYKDDPVDLR LDIERRKHHKERDLKRGKSRSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS							
NLPYNSTSYPNYFOHRTQKEASISWESSLFPA LVQTNCYXYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWFEDTDCSQ FPEENSDNQTCLMPDEYVECESPSHFKCRSGQ CVLASRRCDGQADCDDGDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACYSDL WCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQELLSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLMGTTLHELLVNGQS CESSRSKISLLCTKQDCGRRPAARMIKRILGGR TSRPGRWPWQCSLOSEPSHICGCVLLAKKW VLTVAHCFEGRENAAVWLVVLGDNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRVCLPNFEQWLEPDTYCYTIGW GHMGNKMPFKLQEGEVRISLEHCQSYFDMK TITTRMICAGYESGTYDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIRKQTYOTPLIA TITRMICAGYESGTYDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIRKQTYOTPLTYFIGG NRGFYPWGQVJNRGGYGNYSNWQNYRQAY SPRSSSRSSSSSSSSSRSRSSSSSSSSSSSSSSSSSSS							
LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ PPEENSDNQTCLMPDEVVECSPSHFKCRSQ CVLASRRCDGQADCDDDSDEENGCKERDL WECPSNKQCLKHTUCOGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVICOGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVICOGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVISSLMVHRAATEHH VCADGWQELLSQLACKQMGLGEPSYTKLIQE QEKEPRWLTLHSNWESLMGTTLHELLVNIQGS CESSRSISLLCTKQDCGRAPAARMIKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLLAKKW VLTVAHCFEGRENAAVWKVVLGINNLDPS VFMQTRFVKITILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTIGW GHMGNIKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCPSKVLGPGVYSNVS FYVEWIRQIYINFSKSGSSESSSRSASSRSRSRS SKSRSSRSSRSSRSSRSSRSSRSSRSSRSSRSSRSSRSRSRS			1				
CRĂLCEHSKERCESVLGIVGLQWPEDTDCSQ PPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ CVLASRRCDGQADCDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPCDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSTLMVHRAATEHH VCADGWOELISQLAKGMGLGEPSVTKLIOE QEKEPRWI.TLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMIKRILIGGR TSRFGRWPWQCSLOSPGHICCCVLIAKKW VLTVAHCFEGERNAAVWKVVLGINNLDHPS VFMQTRFVKTILLHPRSAVVDYDISIVSVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTTGW GHMGNKMPFKLQEGEVRISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGFLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS FYFEWIKRQTYQTFLIAN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSSSRSSSRSSRSSRSSRSSRSSRSSRSSRS	1	Í	ľ				-
PPEENSDNOTCLMPDE YVEEC SPSHEK CRSQ CVLASRRCDQ ADDDD SDEENCG CKERDL WECPSNKQ CLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDE WDCVTLSINVNSSSFLMVHRATEHH VCADGWCELLSQLACKQMGLGEPSVTKLLOE QEKEPRWLTHLSNWESSFLMVHRATEHH VCADGWCELLSQLACKQMGLGEPSVTKLLOE QEKEPRWLTHLSNWESSFLMVHRATEHH VCADGWCELLSQLACKQMGLGEPSVTKLLOE QEKEPRWLTHLSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRAARMNKRULGGR TSRPGRWPWQCSLQSEPSGHIGGCVLLAKKW VLTVAHCFEGGENAVWEVVLGINNLDHPS VFMQTKFVKTTILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTITGW GHMGNKMPFKLQGEVRIJSLEHCQSYFDMK TITTRMCAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQTYQTFLLN SYSTEMSSKSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS							
CVLASRRCDGQADCDDDBEDROGCKERDL WECPSNKQCLKHTVICDGFPDCPYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVINSSSFIM/HRAATEHH VCADGWGELSQLACKQMGLGEPSVTKLIOG QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESSRKISLLCTKQDCGRRAARM/KRILIGGR TSRFGRWPWQCSLQSEPSGHIGGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTIGW GHMGNKMPFKLQEGEVRIISLEHCQSYFPMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNISKSGSRSSRSRSRSRSRSRSRSRSRSRSRSRSRSRSRSRS							•
WECPSNKQCLÄHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDL WCDGEADCS DSSDEWDCVTLSNYNSSSFLMVHRAATEHH VCADGWGEILSQLACKQMGLGEPSVTKLIQE QEKEPRWITHLHSNWESLNGTTHHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGR TSRPGRWPWQCSLQSEPSGHIGGCVLIAKKW VLTVAHCFEGGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTIGW GHMONKMPFKLQGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLIN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASKSRSRS FSKSRSRSLSSRSRKRLSSRSRSRSRSASKSRSRS FSKSRSRSLSSRSRKRLSSRSRSRSRSSRSSSSSSSSSSS							,
CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSTLMVHRAATEHH VCADGWQELISQLACKQMGLGEPSVTKLIQE QEKEPRWLITHISNWESLNGTTHELLVNGQS CESRSKISLLCTKQDCGRRPAARMKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWCVUGINNLDHPS VPMQTRFVKTIILHPRYSRAVVDVDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYITGW GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITITRMICAGYJESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGYYSNVS YFVEWIKRQIYIOFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSSSSSSSSSSSSSSSSSSSSSSSS	1		1				
DSSDEWDCVT.SINVNSSSFLMVHRAATEHH VCADGWGELSQLACKQMGLGEPSVTKLIQE QEKEPRWI.TIHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGR TSRPGRWPWQSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VPMQTRFVKTIILHPRYSRAVVDVDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTTGW GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLISWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLIN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSSRSSRSRSRS SSRSRSRSSRSSRSSRSRSRSSSPAHN RERNHPRVQQNRDFRGHNRGYRPFYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGSRSSRSPKRRSPSPRSRSHSNSDKSSS DRSRRSSSSSNISRVESSKRKSAKEKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPFRPSPVPKPSPPLSST SQMGSTLPSGAGYQGSTHQGCPHGSGGLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEGQKTENGKDKEQKQTNTDKEKKE KGSFSDTGLGGDKMSDSFS PKTDSEKPRE SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEFKFNSKVIGANKNQEEE KSGKWGGLVYAPPGKEKQRTTDSEKPRE SSSKEDDEFKPNSKVIGANKNQEEE KSGKWGLVYAPPGKEKQRKTEELEEESPE KSCKWGLVYAPPGKEKQRKTEELEEESPE RSKKEDRGRRSEGGHRGFVPEKNFVTAYK AVQEKSSSPPPRKTSESRKLGAKGGDFTIGKS SFSITTRAQVNVRMDSSTDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIPGHIQSAQSQRSP SELFAQHIVTIVHHVKEHFIGSSGMTLHERFT KYLKRGTEGEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGSKCKKKRRARDSRSSS	l						
QEKEPRWİ.TLHSNWESI.NGTTLHELLVNGÓS CESRSKISLLCTKQDCGRRPAARMINKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCPEGRENAAVWKVVLGINI.DHPS VFRQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTITGW GHMGNKMPFKLQEGEVRIISLEHCQSYPDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLISWGSVCPSKVLGPGVYSNVS YFVEWIKRQIVIOTILIN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSYSAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQVNRGGYGNYSSNWQNYRQAY SPRRGRSRSSSRSSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGNOGDEVKEQTTSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPPSPVKPSPPLSST SQMGSTLPSGAGYQSGTHQGPDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNITDKEKIKE KGSFSDTGLGDGKMKSDSFAFKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPFFMSKVIGANKNQEEL KSGKWEGLVYAPPGKEKQRKTEELEESPPE RSKKEDRGKRSEGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESROKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKEQEFRSIFQHQSAQSQQSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR	İ	•	•				
CESRSKISLLCTKQDCGRPAARMNKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTTGW GHMGNKMPFKLQGGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQJVIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS FSKSRSRSLSRSRKRLSSRSRSRSYSPAHN RERNIPRVYQNDFFRGHNRGYRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWONYRQAY SPRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRSSSSSSSSSNSNSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRES SPALKSPLQSVVVRRSPRPSPYPKPSPLSST SQMGSTLPSGAGYQSGTHQQGFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSCKESAASGGAA YTKRYLEBCKTENGKDKEQKOTNITDKEKKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAAGGKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSSGPHRGFVPEKNFRVTAYK AVQEKSSSPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRNDSFDEDLARPSGLAQER KLCRDLVHSNKKEQEFFSIFQHIQSAQSQRSP SELFAQHIVTIVHIVKEHIFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR		1					VCADGWQEILSQLACKQMGLGEPSVTKLIQE
CESRSKISLLCTKQDCGRPAARMNKRILGGR TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTTGW GHMGNKMPFKLQGGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQJVIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRSRS FSKSRSRSLSRSRKRLSSRSRSRSYSPAHN RERNIPRVYQNDFFRGHNRGYRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWONYRQAY SPRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRSSSSSSSSSNSNSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRES SPALKSPLQSVVVRRSPRPSPYPKPSPLSST SQMGSTLPSGAGYQSGTHQQGFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSCKESAASGGAA YTKRYLEBCKTENGKDKEQKOTNITDKEKKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAAGGKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSSGPHRGFVPEKNFRVTAYK AVQEKSSSPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRNDSFDEDLARPSGLAQER KLCRDLVHSNKKEQEFFSIFQHIQSAQSQRSP SELFAQHIVTIVHIVKEHIFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR		ļ					QEKEPRWLTLHSNWESLNGTTLHELLVNGQS
VLTVAHCFEGRENÁAVWKVVLGINNLDHPS VFMQTRFVKTILHPRYSRAVUDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTIGW GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYBEGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNSKSGSRSSRSSRSSRSSRSSRS KVILSSEMSKTNKSKGGTSSSRSSRSASRSRSRS FSKSRSRSISSRSRSRSRSRSRSSRSSRSSRSSRS FSKSRSRSSRSSRSSRSSRSSRSSRSSRSSRSSRS PSKSRSRSSRSSRSSRSSRSSRSSRSSRSSRSSRS PSKSRSRSRSSRSSRSSRSSRSSRSSRSSRSSRS FSKSRSRSRSSRSRS							
VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE DISETGYVRPVCLPNPEQWLEPDTYCYTIGW GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSASKSRSRS FSKSRSRSRSLSRSRKRLSSRSRSRSRSRSPAHN RERNHPRVQNDFDRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSPRSSHSRNSDKSSS DRSRRSSSSRSSSNISRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSFRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTITDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQINTTVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKGKSRESVDSRDSSHSR ERSAEKTEKTHKGSSKCQKKKHRRADDSSSSS	1						TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW
DISETGYVRPVCLPNPEQWLEPDTYCYITGW GHMGNKMPFKLQGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSRSRSRSRS FSKSRSRSSRSSRSRSRSRSRSRSRSRSRSRSRSRSRSRS	1	ļ					VLTVAHCFEGRENAAVWKVVLGINNLDHPS
GHMGNKMPFKLQEGEVRISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRSS FSKSRSRSRSLSRSRKRLSSRSRSRSRSSPFARRSPSYPAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSSRSSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSSRSSNSHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQITSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQPDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTELLEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SCLFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFFK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS				[VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
TITTRMICAGYESGTVDSCMGDSGGPLVCEK PGGRWTLFGLTSWGSVCPSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGGRSSRSRSASRSRSR FSKSRSRSRSLSRSRKRRLSSRSRSRSRSRSRSRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSRSPAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSWQNYRQAY SPRRGSRSRSPSKRRSPSPRSRSHSRNSDKSSS DRSRRSSSSRSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQITSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQFFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRDISFSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	ŀ						
PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSRSRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSSRSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSRSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTTSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFFTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEHRRIDISPSTTRK HGLAHDEMKSPREPGYKAEGKYKDDPVOLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	1						
996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSKRSKSRSRSS FSKSRSKSLSRSKRKLSSRSRSRSRSRSRSRS FSKSRSRSLSRSRKRLSSRSRSRSRSRSPAHN RERNHPRVYQNRDFRGHINRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGSRSRSPKRRSPSFRSRSHSRNSDKSSS DRSRRSSSRSSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTTSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTINTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPRKTSESRDKLGAKGDFFTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHIVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKSPEHRRIDISPSTERK HGLAHDEMKSPREPGYKAEGKYKDDPVOLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	İ						
996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSRSRSRSRSFSPAHN RERNHPRVYQNRDFRGHINRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTTSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDDFPKKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRSSVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	i .			}			
FSKSRSRSRSRSRSRSRSRSSYSPAHN RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRSSSSRSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	006	00.45		22.22			
RERNHPRVYQNRDFRGHNRGYRRPYYFRGR NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSRSHSRNSDKSSS DRSRRSSSRSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFKK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	996	2346	A	8392	199	3085	
NRGFYPWGQYNRGGYGNYRSNWQNYRQAY SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSSRSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPPPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEESFPE RSKKEDRGKRSEGGHRGFVYEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							· · · · · · · · · · · · · · · · · · ·
SPRRGRSRSPKRRSPSPRSRSHSRNSDKSSS DRSRRSSSSRSSNHSR VESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTPSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYJLRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	ļ			ĺ			
DRSRRSSSRSSNHSRVESSKRKSAKEKKSSS KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS			•				
KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
ASESSKPWPDATYGTGSASRASAVSELSPRER SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFTT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS					:		
SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYYLKGGTEQEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							YTKRYLEEOKTENGKDKEOKOTNTDKEKIKE
SQSPKRYKLRDDFEKKMADFHKEEMDDQDK DKAKGRKESEFDDEPKFMSKVIGANKNQEEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
DKAKGRKESEFDDEPKFMSKVIGANKNQÈEE KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYJLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS			,				
KSGKWEGLVYAPPGKEKQRKTEELEEESFPE RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
SFSITREAQVNVRMDSFDEDLARPSGLLAQER KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS
SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS					İ		SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT
LDIERRKKHKERDLKRGKSRESVDSRDSSHSR ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS							
ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS	1					. :	
				•			
SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL							
							SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL

NO: of NO: of hod ID N	O: Predicted beginning nucleotide	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, E=Bhomalologies C=Charine H=Histidiae
nucl- peptide in USS seq- uence 09/4	location	location corresponding to last amino	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence 914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
	sequence		nucleotide insertion
			GTKDFVGPSERGGGRARGTFQFRARGRGWG
	1	-	RGNYSGNNNNNSNNDFQKRNREEEWDPEYT
	· ·		PKSKKYYLHDDREGEGSDKWVSRGRGRGAF
			PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE
997 2347 A 8398	202	552	DDESGTENREEKDNIQPTTE CPALGGRODLQGTRLLWAHDSGVGGOKAKS
		552	KQENLESLEATGREEEGGQGPPVTTKGVLLA
			LLMAGLALQPGTALLCYSCKAQVSNEDCLQ
			VENCTQLGEQCWTARIREWGDDSRQA
998 2348 A 8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN
1 1. 1			MSDPRRPNKVLRYKPPPSECNPALDDPTPDY MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI
		ļ	SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ
			NPOPMTPPW
999 2349 A 8401	93	1126	ASASHITSGHLRCPPGSEGVGTMARCFSLVLL
]]]	ļ	1	LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL
1			VSKKANQQLNFTEAKEACRLLGLSLAGKDQ
	<u>'</u>		VETALKASFETCSYGWVGDGFVVISRISPNPK CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW
			TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS
			VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE
			VFMETSTMSTETEPFVENKAAFKNEAAGFGG
	1	[VPTALLVLALLFFGAAAGLGFCYVKRYVKAF
			PFTNKNQQKEMIETKVVKEEKANDSNPNEES KKTDKNPEESKSPSKTTMRCLEAEV
1000 2350 A 8406	2	777	KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT
	_	[]	AAIFTADGNMISASTLMDILLMNDFKLVINKI
			AYDVQCPKREKPSNEHTAEMEHMKSLVHRL
			FTILHLEESQKKREHHLLEKIDHLKEQLQPLE
1 1 1			QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA LAWLTWWVYSWDIMEPVTYFITFANSMVFF
			AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ
			HFDVQQYNKLKEDLAKAKESLKQARHSLCL
			QMQVEELNEKN
1001 2351 A 8410	1400	264	VGFWERPLRSSRWFRRSLRRWEMLARAARG
			TGALLLRGSLLASGRAPRRASSGLPRNTVVLF
]	}		VPQQEAWVVERMGRFHRILEPGLNILIPVLDR IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV
			LYLRIMDPYKASYGVEDPEYAVTQLAQTTM
			RSELGKLSLDKVFRERESLNASIVDAINQAAD
	1		CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE
			RRKRATVLESEGTRESAINVAEGKKQAQILAS
		.*	EAEKAEQINQAAGEASAVLAKAKAKAEAIRI LAAALTQHNGDAAASLTVAEQYVSAFSKLA
	}		KDSNTILLPSNPGDVTSMVAQAMGVYGALT
	<u> </u>		KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR
1000			VKMS
1002 2352 A 8421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY
			YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL
·			AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC
			PKCTQIINGSVDVDTEDRQKRKPESDGRTAK
	_ [ALRSLQFTNPGRQTEFAPETGKREKRRLTKN
	1		ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC
	[DCLDEDCLGCFYACPACGSTKCGAECRCDRK
1003 2353 A 8427	3	1416	WLYEQIEIEGGEIIHNKHAG TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR
		1410	CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT
			MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence	sequence	/=possible nucleotide deletion, \=possible nucleotide insertion
						SLDTENIDEILNNADVALVNFYADWCRFSQM LHPIFEEASDVIKEEFPNENQVVFARVDCDQH SDIAQRYRISKYPTLKLFRNGMMMKREYRGQ RSVKALADYIROOKSDPIOEIRDLAEITTLDRS
						KRNIIGYFEQKDSDNYRVFERVANILHDDCAF LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY LGAMTNFDVTYNWIQDKCVPLVREITFENGE
•						ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL ISEKGTINFLHADCDKFRHPLLHIQKTPADCP VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL
1004	2354	A	8432	910	387	HSGKLHREFHHGPDPTDTAPGEQAQDVASSP PESSFQKLAPSEYRYTLLRDRDEL
1004	2534,	A	0432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM ACAAARSPADQDRFICIYPAYLNNKKTIAEGR RIPISKAVENPTATEIQDVCSAVGLNVFLEKN KMYSREWNRDVQYRGRVRVQLKQEDGSLC
1005	2355	Α	8453	90	530	LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA DQSLQQGEGSKKGKGKKKK QSHETKMQSGTHWRVLGLCLLSVGVWGOD
						GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF SELEQSGYYVCYPRGSKPEDANFYLYLRARG NPGLQNRYHRLFREDHSKGHSQ
1006	2356	A	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC QLCIFN
1007	2357	A	8459	43	553	GAGAGGDWAAMDKLKKVLSGODTEDRSGL SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA FWWHNKGLALIFCILQSLALTWYSLSFIPFAR DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS DPRWGCVGPSMPTSTCLPGAVEASTTKASLP KCPVDSSLPTPEACFL
1009	2359	A	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN
				·		FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL TKLLVHSSLVGSILSALSALVGFIILSVKQATL NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT
1010	2360	A	8468	2	473	HDCGYEELLTS KYRYRRPYPVMRKICQVGPAGLAFILNISPVA HRVALCHLAGCQEQAAWYHTLQILFFLVSAY FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL
1011	2361	A	8478	5	409	SFFFLAACSAATAALLRHKVKARLTKKDS TELSQLEKAHPPADMGRRKSKRKPPPKKKMT
						GTLETQFTCPFCNHEKSCDVKMDRARNTGVI SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC RAGFQCQ
1012	2362	Α .	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM RMKYGGQEFWADLNAMNVYETTEFDQLRR LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

NO. of No. of not	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Deptide code Sequence Seque							
Sode Sequence Sod				1			
uence 09/495 corresponding to find first and the state of peptide sequence foreigned peptide se			1	1			
union gn first samino aeid residue of peptide pe	1						· · · · · · · · · · · · · · · · · · ·
mino add residue of peptide requence per per per per per per per per per pe		}]				
residue of peptide sequence	1		ŀ			of peptide	
Peptide			İ	İ			
		!					
SVIRLIEEANSRÜLKEVRPAMWNINTTLINS FERREKREHERSCHLEPYLQTIGGOPTOAD PPLEATSSQÜLCPDGVTSANFTYETWVYMHP SQDFIGVPVSAENES KRIVINLIFIKTYPEFKYR LQURVONOPI WEKYKRKEYMRKAMFOR DEINERRIH-HGTSQUVDGIGKENHPERVCG KHATMFOGGSYFAKKASYSINTSKKSSKOV HMFLAKVUTGRYTMGSHÖMRERPPVVPGS VYTSDLNDSCVDNFFEPQIFVIRNDDGSYPYFVI QVIEVSNTVSI STRIGWWFRQPVLYTQSAILVPVRTKKRTTY PTOPKFTREKFMGHARKAGVIPPEKSDRS HHACTAGIFDA VVPFEGDARISSLSKEGLIER TERNKKMASQVSIRRIKOTDANFKKEPT PTOPKFTREKFMGHARKAGVIPPEKSDRS HHACTAGIFDA VVPFEGDARISSLSKEGLIER TERNKKMASQVSIRRIKOTDANFKKEPT PTOPKFTREKFMGHARKAGVIPPEKSDRS HHACTAGIFDA VVPFEGDARISSLSKEGLIER TERNKKMASQVSIRRIKOTDANFKKEDFE KAKDIFIEGSPLY AQLMYTVYTTHSLEIMYSIRTAVYVICHY AQLMYTVYTTHSLEIMYSIRTAVYVICHY AQLMYTVYTTHSKUEDHYSTRATVYVICHY AQLMYTVYTTHSWISCHMYSIRTAVYVICHY AQLMYTVYTTHSWISCHMYSIRTAVYVICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY AQLMYTVYTTHSWYSTRAYVYICHY ALLLEPPDORGARIVANGAGVITAFORYVICH AND AND AND AND AND AND AND AND AND AND	1						
FFRREIKRRPLERSCEILLPYLQTILGGYPTQAB			_				
PPLEATSSQUICPDOVTSANFYPETWYTHEN SQDFIGPYPYSADKSYRINYALFIKTYPEFKYR LQLRVONOPI_WEKYRKKEYMNRKMFG DRINDERHL-PHGTSQDVVDGIGKHNFDPRVCG KHATMFGGGSYFAKKASYSHNFSKCSSKG KHATMFGGGSYFAKKASYSHNFSKCSSKG KHATMFGGGSYFAKKASYSHNFSKCSSKG VTSDLYDSCVDNFFEPQIFVIFNDQSYPYFVI QYEVSNTVSI SER_GWRFQPVLYTOSAAVPAYTIKKRFTP PYCOPKEYSTVSI SER_GWRFQPVLYTOSAAVPAYTIKKRFTP PYCOPKEYSTEKEFMOHARKAGI_VIPPEKSDRS HLACTAGIFDAYVPFEDDAISSLSKEGIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSFL YTRIGYVYICITYAQLMYTYYTIKTAYVYICILY AQLMYTYYYTHASCHMYSIKTAYVYICILY AQLMYTYYYTHALCHMYSIKTAYVYICILY AQLMYTYYYTHALCHMYSIKTAYVYICILY AQLMYTYYYTHALCHMYSIKTAYVYICILY AQLMYTYYYTHALCHMYSIKTAYVYICILY AQLMYTYYYTHASSLVOPLICY TLSHPSYLKOPLYTYYYTHASCHMYSIKTAY AVLTACHPSYNYYCSTOPLIC AVLTACHPSYNYCSTOPLIC AVLTACHPSYNYCSTOPLIC AQLACH AND AND AND AND AND AND AND AND AND AND	1				ļ		
SQDFIQVPV\$AEDKSYRIVPLIFIKTVPEFKY	1		1			i	
ILQURVONOFILWEKYRKKEYMARKMENG	Ì			ļ		ĺ	
DRINNERH_FIGTSQDVVDGICKENTPDPRVQC KIAITMEGGGSTVAKKASYSHRPSKESSKGV HPMFLAKVLTGRYTMGSHGMRPPPVNPGG VISDLYDSCVDNFFEQIFVIRDDQSYPYFVI QYEEVSNTVSI DIVERSON DIVERSON DIVERSON DIVERSON							, ,
KHATMFGQGSYFAKKASYSHNESKSSKGV HEMFLAKVLTGRYTMGSHGMRPPPVNPGG		ĺ		1		[
HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS	J	1]			ļ	
1013 2363 A 8488 2 517	Ī	İ	ŀ	ļ			l
1013 2363 A 8488 2 317	1		1	l			
1013	1	{	i				
SRELGWWERQPYLYTOSAAIYPVERTKEFTP PIYQPKFKTEKEFMQHARKAG VIPPEKSDRS IBLACTAGIFDAYYPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIPIEGSH.Y 1014 2364 A 8501 363 17 YIRTGYVYICITYAQLMYTYYRTAYVYICITY AQLMYTYVYTHSLCHMYSIRTDYVYICITY AQLMYTYVYTHSLCHMYSIRTDYVYICITY AQLMYTYVYTHSLCHMYSIRTDYVYICITY AQLMYTYVYTHSVMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSF YSWYGSARLFRFRVPPD AVLLRWLLQVSRESGAACTDABITHYRSGA PPVINPLGTSPPDDTAVQPSQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKELKG LAFTCAYVFQFELLVTRVGVPTLSIMEPDVPLPQ TILSHPSYLKVFVPDYTREILLELRDCVSNGS LGCPVRLTVGPVTLPSNFGVCVLTCTGAPWPC RLLLFSPPWDRWLQVTAESLVGPLGTVAPSA VAALTACRFRSVTIQPLLGSSQNGSFASSGI LSPSPPHQDLGRGSGVDSSPFGLTNYPVTRED MDVVSVHFQFLDRVSVRVCSDTPSVMRLRL NTGMDSGSSTISLRANKTEMRNETVVVACV NAASPFLGFINTSLNCTTAFFQGYPLSLAWSR RANLIHYPTDTNWYLSLQMCPSALADEQ AVVHVETTI,YLPCLNDGGPYGQCLLRHS YLVASCSCKAGWRGVSCTDNSTAGTVAQQR AATLLLTLSNJMFLAPIASVVRFFLVEASVY AYTMFFSTFYHACDQPGBAVLCLSVDTLQY CDPLGSGAAIWVTILCMASVVRFFLVEASVY AYTMFFSTFYHACDQPGBAVLCLSVDTLQY CDPLGSGAAIWVTILCMASVQRWFFLVFLDG VSMASVGIATYTSMMTSDNYYYTHSIWHILL AGSAALLIPPPDQFAEPWACSQKFPCHYQIC KNOREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRFGHRYCKMAPAK KGGEKKGRANIPVTRIVNGVRWFFLIPG VSMASVGIATYTSMMTSDNYYYTHSIWHILL AGSAALLIPPPDQFAEPWACSQKFPCHYQIC KNOREELYAVT KNOREELYAVT KNOREELYAVT KWYPSGPVRFGRFYYKLPAGHRRCRMAPAK KGGEKKGRANIPVTRIVRSPRKNEDEDSP NKLYTLVTVVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTFASADMAWTKYQLFLAGLMINTTGSINT LAKWADNAFMAEGGGSKEHSPQHFTLQAV GMFLGEFSCLAAFYLLRCRAAGGSSVDPQ QPPFNILFLPALCDMTGISLMYVALNMTSA SSFQMLRGAVITICRFALVSGWL GILATIAGLVVVGLADLLSKHDSOHKLSEVIT GDLLIIMAQIVALQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLVFWYYTRAGSFSGNP RGTLEDALDAFGQVGQQPLIAVALLGNISSIA	1013	2262		0400	2	517	
PIYOPKFKTEKEFMOHARKAGLVIPPEKSDRS HILACTAGIFDAYVPPEGARRSISLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDPPE KAKDIFIEGSPLY TERMKKTMASQVSIRRIKDYDANFKIKDPPE KAKDIFIEGSPLY AQLMYTYVLYTHASCHMYSIRTAYVYICITY AQLMYTYVLYTHASCHMYSIRTAYVYICITY AQLMYTYVLYTHASCHMYSIRTDYVYICILY AQLMYTYVFYTHASCHMYSIRTDYVYICILY AQLMYTYVFYTHSCHMYSIRTAYVYICITY AQLMYTYVFYTHSCHMYSIRTDYVYICILY AQLMYTYVFYTHSCHMYSIRTDYVYICILY AQLMYTYVFYTHSCHMYSIRTDYVYICILY AQLMYTYVFYTHSCHMYSIRTDYVYICILY AQLMYTYVFYTHSCHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE NSSEHFSQAPQRLSF YSWYGSARLFRFRVPPD AVLRWILLQVSRESGAACTDAEITVHFRSGA PVINPLGTSFPDDTAVQPSTQVGPSIVTPRS NASVNVSHPAPGDWFVAAHLPPSSQKELKG LAFTCAYVFQPELLVTRVVEISIMEPDVPLPQ TILSHPSYLKVFVPDYTTELLLERDCVSNGS LGCPVRLTVGPVTLPSNFQVLTCTGAPWPC RILLIPSPWDRWLQVTAEELVGPTAPSQ LGPVRLTVGPVTLPSSQNGSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNTVYTRED MDVVSVHFQPLDRVSVRGVSTPSVMCTLL NTGMDSGGSLTISLRANKTEMRNETVVVACV NASSPLGPTNSTLACTTAFFQGYPJLSAWSR RANLIIFYPETDNWYJSLQLMCFENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLRRHS YLVASCSCKAGWRGWSCDNSTAQTVAQQR AATILLTLSNLMFLAPIAVSVRRFTVLVASVY AYTMFFSTFYHACDQPGGAVLCILSVDTLQV CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLAFVIM ASSMVAYRGHRQCVPTSVMCWAFVLLPG VSMASVGIAITTSMMTSDNYYTISWHILL AGSAALLIPPPDQPAEPWACSQKFPCHYQIC KNDRELYAVT KNAWAKGGNFNYFRIRVSKRNFEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN KNAWAKGGNNYFIRVSLSKRNFEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN LERTPASADMAWTKYQLFLAGIMLVTGINT LSAKWADNFMAEOGGSKHSTQHFTQAV GMFLGEFSCLAAFYLLRCRAAGGSDSVDPQ QPPNPLLFLPALDDMTGJLSMYVALNMTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GILATLGCLVFQLACUSCHDAVSKINTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GILATLGCLVFQLACUSCHDAVSKINTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GILATLGCLVFQLACUSCHDAVSKINTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GILATLGCLVFQLACUSCHDAVSKINTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GILATLGCLVFQLACUSCHDAVSKINTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GILATLGCLVFQLACUSCHDAVSKINTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GRICATAGLOVQCLACUSCHDAVSKINTSA SSFQMLRGAVILTGLFSVAFLGRULSQWL GRICATAGL	1013	2505	^	0400	 	317	
IHLACTAGIFDAYVPPEGDARISSLSKEGLIER TERMKKIMASQVSRRIKDYDANFKIKDFPE KAKDIFIEGSPLY TERMKKIMASQVSRRIKDYDANFKIKDFPE KAKDIFIEGSPLY TERMKKIMASQVSRRIKDYDANFKIKDFPE KAKDIFIEGSPLY TERMKKIMASQVSRRIKDYDANFKIKDFPE KAKDIFIEGSPLY TERMKKIMASQVSRRIKDYDANFKIKDFPE KAKDIFIEGSPLY TERMKKIMASQVSRRIKDYDANFKIKDFPE			1				
TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY AQLMYTYVIYTHSLCHMYSIRTAYVYICILY AQLMYTYVFYTHRICHIMYSIRTDYVYICILY AQLMYTYVFYTHSLCHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHSQAPQRLSFYSWYGGARLFRFRVPPD AVLLRWLLQVSRESGAACTDAEITVHFRSGA PPVNRLGTSFPDDTAVQPSRQVVPLSTIPRS NASVNVSHPAPGDWVAAHLPPSSQKELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLISHPSYLKVFVPDVTRELLLEIRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPVPC RILLPSPPWDRWLQVTAESLVGPLGTVAPSA VAALTACRPRSVTIQPLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSFFCLTNYPVTRED MDVVSVHFQPLDRVSVCSDTPSVMRIRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPLGFNTSINCTTAFQQVPLSLAWSR RANLIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETILYLVPCLNDCGFYGQCLLLRRHS YLYASCSCKAGWRGWSCTINDSTAQTVAQQR AATLLITLSNLMFLAPIAVSVRRFFLVEASSVY AYTMFFSTFYHACDQPGAVLCLLSYDTUQY CDFLGSGAAIWVTLCMARLKTVLKYVLFLL GTLVIAMSLQLDRGMWNNLGCFLAFVIN ASSMASVGIAYTSMMTSDNYYTHSIWHILL AGSAALLIPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRCCRMAPAK KGGEKKGRSAINEVVTREYTINHKRHGVG FKKRAPRALKEIRKFAMKEMGTDVRDTRL NKAVWAKGRINNYPKIRVRLSRRNEDEDDSP NKLYTLVTYVPVTTFKNLQTVNVDEN LSAKWADNFMABGCGSKEHSPQHFFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPG QPFNPLLFLPPALCDMTGTSLMYVALINMTSA SSFQMLRGAVITFGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLALDKHSDKHLSEVIT GDLLIIMAQUIVALQMVLEEKFVYKHNVHPLR AVGTEGLFGFVLISLLVPMYTYPAGSFSCIN FGTLEDALDAFCQVQQCPLLAVALLGINISSIA	1		1			}	l
1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVYICIIYAQLMYTYYIRTAYVYICIIYAQLMYTYYIRTAYVYICIIYAQLMYTYYIRTAYVYICIIYAQLMYTYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYYIRTAYVYICIIYAQLMYTYYYYITHIYAGAALPRISTAYVYYICIIYAQLMYTYYYIRTAYVYICIIYAQLMYTYYYIRTAYAGAALPRISTAYAQQRAQAAATAACRPRSYTIQPLQSSQOQSFNASSGLL LSPSPDHODLGRSGROYDRSFFCLTNYPYTREDMDVSVHFQPLQRSVQRSFNASSGL LSPSPDHODLGRSGROYDRSFFCLTNYPYTREDMDVSVHFQPLDRVSVRVCSDTPSVMRIRI.NTGMDSGSLTISIRANKTEMRNETVVVACVNAASPFLGFNTSINCTITAFFQGYPLSISAWSR RANLIPYPETDNWYLSIQLMCPENAEDCEQ AVYHVETTLYLVPCLINCGPYGQCLLIRRISTYLYAGSCSCKAGWRGWSCTDNSTAQTVAQQR AATLLITISNLMFLAPIAVSVYRFFLVEASVY AYTMFFSTFYHACDQPGBAVLCLISYDTLQY CDFLGSGAATWVTLCMARIKTVLXYVILLI GTT.VIAMSLQLDRRGMWNMLGPCLFAFVIMAASMAYRCGHRQCYPTSWQRWAFYLLPG VSMASVGIAYTSMMTSDNYYYTHSIWHILL AGSAALLIPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYFSGPVRIPGRFYYKLPAGHRCRMAPAK KGGEKKKGRSANEVVTREYTINHKRHGYOG FKKRAPRALKEIRKFAMKEMGTPDVRIDTIL NKAVWAKGRNVPYRRVRLSRKRNEDEDSP NKLYTLYYVPUTTFKNLQTVNODEN LSAKWAGRRNVPYRRVRLSRKRNEDEDSP NKLYTLYYVPUTTFKNLQTVNODEN LSAKWADNFMAEGCGGSKEHSSQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVPQ QPFNPLLFLAGILMUTGSINTY LSAKWADNFMAEGCGGSKEHSSQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSASSFQMLRGAVITFGLFSVAFLORRIVLSQWL GILLIMAQUVALQMVLEEKFYYKHNVHPLR AVGTECLFFVIKANVHPLR AVGTECLF			1	1			1
1014 2364 A 8501 363 17			l				
AQLMYTYVLYTHSLCHMYSRTAYVYICHY AQIMYTYVFYTHSVMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSAACTDAEITYHFRSGA PPVNPLGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAAGDWFVAAILPPSQKELKG LAPTCAYVFQPELLVTRVEISIMEPDVPLPQ TLLSHPSYLKVFYPDYTRELLLERDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSONGSFNASSGL LSPSPDHQDLGRSGRVDRSFFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGSLTISLRANKTEMRNETVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIPYPETDNWYLSLQLMCFPNAEDCEQ AVYHVETTLYLVPCLDNCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTIDNSTAQTVAQQR AATLLLTSNLMFLAPAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTLCMARLKTVLKVYLFLL GTLVTAMSQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRQCYPTSWQWAFYLLPG VSMASVGIAYTSMMTSDNYYYHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYFSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSADNEVVTREYTINHKRHIGYO FKKRAPRALKEIRKFAMKEMGTPDVRTIDTL NKAVWAKGRNNPYRRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNUDEN LSARWAGRNAPTYRRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNUDEN LSARWAGRANFVTRAFTRANGTPDVTRIDTL NKAVWAKGRNNPYRRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNUDEN LSARWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVHTFGLFSVAFLORRIVLSQWL GILATIAGLVVGLALAGLKSHSONIKLSEVIT GDLLIMAGUVVALADMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLVVPMYYPAGSFSGNP RGTLEDALDAFCQVGQQPLLAVALLGNISSIA	1017	226	<u> </u>	0.56=		L	
AQİMYTYVFYTHRLCİHMYSIRTDYVYİCILY AQİMYTYVFYTHSYMSDE	1014	2364	A	8201	363	17	3
AQLMYTYPYHSYMSDE		1	i				
1015 2365 A 8504 3 2190 NSSEHFSQAPQRI.SFYSWYGSARI.FRREVPPD AVLLRWILLQVSRESGAACTDAEITVHFRSGA PVVINPI.GTS:PPDDTAVQPSTCYGVGVPLSTTIPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFOPELLUTREVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPYRLTVGPVTLPSNFGKVLTCTGAFWPC RILLISPSPWDRWQUYTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSFFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNISLNCTTAFFGGYPLSLSAWSR RANLIIPYPETDNWYLSQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRIS YLYASCSCKAGMGWSCTDNSTAQTVAQQR AATLLLT.SNLMFLAPIAVSVRFFI.VEASVY AYTMFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTLCMARIKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAVIM ASMWAYCGHROCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLIPPPDQPAEPWACSQKFPCHYQIC KNDREEL YAVT AGSAALLIPPPDQPAEPWACSQKFPCHYQIC KNDREEL YAVT AGSAALLIPPTOQFT SWASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLIPPTOQFT AGSAALTATATATATATATATATATATATATATATATATATA			! .				1
AVLLRWLQVSRESGAÁCTDAEITVHERSGA PPVINPLIGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAPSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTENRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCFENAEDCEQ AVVHVETTLYLVPCLNDGGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCLSYDTLQY CDFLGSGAAIWVTLCMARIKTVLKYVLFLL GTLVTAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPDAAEPWACSQKPPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRPGRFYYKLPAGHRRCRMAPAK KGGEKKGRSAANEVVTREYTTNIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNYPYRIRVLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGIMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFNELFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATTAGLVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGVILSLLLVPMYYPAGSPSGNP RGTLEDALDAFCQVGQQFLLAVALLGNISSIA			<u> </u>				
PPVINPLGTSPPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFQPELLVTREVEISIMEPDVPLPQ TILLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLSPPWDRWLQVTAESLVGPLGTVAPSA VAALTACRPSSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQYPPLSLSAWSR RANLIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRIS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMPLAPJAVSVRFFLVEASSY AYTMFPSTPYHACDQPGEAVLCLLSYDTLQY CDFLGSGAAIWVTLCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAPVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAYTSMMTSDNYYTHSIWHILL AGSAALLLPPPDQPAEFWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGRNVPYRIRVRLSKRNEDEDSP NKLYTLVTVYPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMTVALINMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVYGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGVILSLLLVPMYYPAGSFSGNP RGTLEDALDAFCQVGQQPLLAVALLGNISSIA	1015	2365	Α	8504	3	2190	
NASYNVSHPAPGGWFVÄAHLPPSSQKIELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAPSA VAALTACRPSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPIGFNTSINCTTAFFQGYPLSLSAWSR RANLIPYPETDNYVLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTIDNSTAQTVAQQR AATALLTLSNLMFLAPIAVSVRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDPLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMYNMLGPCLFAPVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAYTSMMTDNYYYTHSIWHILL AGSAALLIPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSANEVVTREYTINHKRIHGVG FKKRAPRALKERKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRRVLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGIMLVTGSINT LSAKWADNFMAEGCGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFYYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDARCQVGQQPLIAVALLGNISSIA							AVLLRWLLQVSRESGAACTDAEITVHFRSGA
LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TILSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAFWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTiQPLLQSSQNQSFNASSGL LSFSPDHQDLGRSGRVDRSFFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETITLYLPVELDDCGPYGQCLLLRHIS YLYASCSCKAGWRGWSCTDDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCLISYDTLQY CDFLGSGAALVVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAPVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIATVSTMMTSDNTYYTHSIWHILL AGSAALLLPPDQPAEPWACSQKFPCHYQIC KNDRELLYAVT KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAYWAKGRNVPYRRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPOHPPLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIJFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFYYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYPRAGSFSGNP RGTLEDALDARCQVGQQPLIAVALLGNISSIA			İ]			PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS
TLLSHPSYLKVFVPDYTRELLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVSVHFQPLDRVSVVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRFFLVEASVY AYTMFFSTTYHACDQPGEAVLCLLSYDTLQY CDFLGSGAAIWVTLCMARLKTVLKYVLFLL GTLVIAMSLQLDRGMWNMLGPCLFAFVIM ASSMVAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDPAEFWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSANEVVTREYTINHKRHIGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRRVRLSKKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTMASDMAWTKYQLFLAGILMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSSQMLRGAVIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPUISLLLVPMYYPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1	ł	}	i i			NASVNVSHPAPGDWFVAAHLPPSSQKIELKG
LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAPSA VAALTACRPRSVTQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETLYLVPCLNDCGPYGQCLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTTPVHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTLCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWMMLGPCLFAFVIM ASMWAYRCGHRQCYPTSWQRWAFYLLPG VSMASVGIAYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT KWYPSGPVRPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSANEVVTREYTINHKRHIGVG FKKAPRALKEIKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN LSRTRAGRENVPYRRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN LSRTRAGRENVPYRRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN LSRTRAGRENVPRGFTSLAMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPRNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLVPMYYPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA				ļ			LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ
RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLISNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKRGRSANEVVTREYTINHKRHGVG FKKRAPALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIVSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN NKLYTLVTYVPVTTFKNLQTVNVDEN LERTPASADMAWTKYQLFLAGILMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHFFLQAV GMFLGFFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIETGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIMMQIIVAJQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS
VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSFFCLTNYPVTRED MDVVSVYHPQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIPYPETDNWYLSLQLMCPENAEDCEQ AVHVETTLYLVPCLNDCGPYGQCLLLRHIS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTPYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARIKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIATYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKGRSANEVVTREYTINHKRHIGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGRNVYPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLIFGFVILSLLVPMYYPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	ĺ						
LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEFWACSQKFFCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIFGRFYYKLFAGHRRCRMAPAK KGGEKKKGRSANEVVTREYTINIHKRHIGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LETTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQUIVAIQMVLEEKFYYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYPAGSFSGIP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	İ						RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA
LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEFWACSQKFFCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIFGRFYYKLFAGHRRCRMAPAK KGGEKKKGRSANEVVTREYTINIHKRHIGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LETTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQUIVAIQMVLEEKFYYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYPAGSFSGIP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1						VAALTACRPRSVTIQPLLQSSQNQSFNASSGL
NTGMDSGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRFFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSANNEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SFFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA			l				LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED
NTGMDSGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRFFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSANNEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SFFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	Ĭ		ĺ	1			MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL
RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRHS YLYASCSCKAGWGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGIMLVTGSINT LSAK WADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQIVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFSSTPYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKGRSANIEVVTREYTINHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNPYPRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGFFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYPPAGSPSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1]			NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR
AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQIVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFSSTPYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKGRSANIEVVTREYTINHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNPYPRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGFFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYPPAGSPSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1						RANLIIPYPETDNWYLSLQLMCPENAEDCEQ
YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLITLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSANEVVTREYTINHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLVPMYYPPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
AATLLITLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIMAQIIVAIQMVLEEKFYYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA]	
AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1]			
CDFLGSGAAIWYTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMILVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GOLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	'						
GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMILVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA					1		·
ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA			l				COM LEGISLATION OF DEPOS PROPERTY OF CONTRACTOR OF CONTRAC
VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAK WADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAK WADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
KNDREELYAVT 1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAK WADNFMAEGCGGSKEHSFQHFPLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
1016 2366 A 8511 1 453 KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	}						
KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1016	2366	A	8511		453	
FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1.0.0	₩00	4.	0511	^	133	I I
NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMILVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GOLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	{ i	'					
NKLYTLVTYVPVTTFKNLQTVNVDEN 1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSPQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
1017 2367 A 8513 54 1196 LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGPVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1010	50/5	 _	0570		1104	
GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1017	/٥٤	A	8212	J 4	1170	
QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							, , ,
SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA							
GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA	1				l		
GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA						ļ	
AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA						ĺ	
RGTLEDALDAFCQVGQQPLIAVALLGNISSIA						1	
	{				ł	ł	1
	·	'}			İ		RGTLEDALDAFCQVGQQPLIAVALLGNISSIA
							FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA S
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLAVLL AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEAAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLLSLTSTLYLRLKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV MAVAVSVGLLLLVVAVFYCVRKKGGPCCRQ RREKGAP
1020	2370	A	8530	2	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	A	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TAEALQL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAIVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYTTLCIVFLM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	}	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	сопеspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
			1	sequence	ļ.	nucleotide insertion
		├──		30quenoc	 	TCKPPLYMGPEYIKYFNDKTIDEELERDKRVT
i	Í	1	í I		ĺ	
	i					WIVEFFANWSNDCQSFAPIYADLSLKYNCTG
1						LNFGKVDVGRYTDVSTRYKVSTSPLTKQLPT
		İ				LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN
1	1	}				VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS
	<u> </u>					TPTTVSDGENKKDK
1025	2375	Α	8546	2194	1707	TVSFHKTMASLKCSTVVCVICLEKPKYRCPA
1	i	ļ				CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS
1					ł	ALPTKTVKPVENKDDDDSIADFLNSDEEEDR
1		ł	i		ł	VSLQNLKNLGESATLRSLLLNPHLRQLMVNL
						DQGEDKAKLMRAYMQEPLFVEFADCCLGIV
1			1			EPSQNEES
1026	2376	Α	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS
1	ļ	1]			YAWANFTILALGVWAVAQRDSIDAISMFLGG
1	1					LLATIFLDIVHISIFYPRVSLTDTGRFGVGMAIL
1					Ì	SLLLKPLSCCFVYHMYRERGGELLVHTGFLG
1	1	i		,		SSODRSAYOTIDSAEAPADPFAVPEGRSODAR
1	į					GY
1027	2377	A	8557	1	340	DFLGPASPQEEGGSESSTMTELETAMGMIIDV
1027	~ ′′	^	0557	*	1 340	FSRYSGSEGSTQTLTKGELKVLMEKELPGFLO
1	[SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF
1		1				VAAITSACHKYFEKAGLK
1028	2378	<u> </u>	8569	20	963	
1028	23/8	A	8269	20	963	KMAATLGPLGSWQQWRRCLSARDGSRRLLL
						LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP
i						YQGEAPRPCFLRDWELQVHFKIHGQGKKNL
		1				HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG
l]	Į				VFVDTYPNEEKQQERVFPYISAMVNNGSLSY
1		!				DHERDGRPTELGGCTAIVRNLHYDTFLVIRY
		•	1			VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
1						YYFGTSSITGDLSDNHDVISLKLFELTVERTPE
i			i l			EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL
1	}	l	1			ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK
		ĺ	1 1			RFY
1029	2379	Α	8572	1	578	AAAASHRSRARSRPRRVSSGPAPRRAQSSAG
						RVASGLDSAPLCTMARALCRLPRRGLWLLLA
]	J			HHLFMTTACQEANYGALLRELCLTQFQVDM
			[EAVGETLWCDWGRTIRSYRELADCTWHMAE
			'			KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR
						AVRDPPGSILYPFIVVPITVTLLVTALVVWQS
						KRTEGIV
1030	2380	A	8574	1352	372	DSSTVKGGSESRHLCLIPDLKGKARTREASSG
1.000	2000		55/4	1006	3,2	SRTCGRRTSLCTSAKSSWTYRSGRLSWOSIKG
1						THLTITOALROPLHRAPLLPGOLCWSPRPLEK
1.						
						NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP
						SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL
						AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY
						VNRLFLNWTEGQESGFLRISNLRKEDQSVYF
					İ	CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT
1				ļ		TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT
1				!		AIRVALAVAVLKTVILGLLCLLLLWWRRRKG
				i		SRAPSSDF
1031	2381	A	8580	905	340	RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL
1				-		AFDDFQESCAMMWQKYAGSRRSMPLGARIL
				l		FHGVFYAGGFAIVYYLIOKFHSRALYYKLAV
						EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI
						VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW
[]						HLDEVFLELKDGQQIPVFKLSGENGDEVKKE
1032	2382	A	8593	2558	961	RRPRLLPGAEPCEPRVGPRRADMGCSAKAR
1032	2002	^	0.75	ەدىي	201	
						WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD VMNPSEILKGEKPQVRERGPYVYREFRHKSNI
1033	2383	A	8595	595	767	TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV MPNILVLGAAVMMENKPMTLKLIMTLAFTTL GERAFMNRTVGEIMWGYKDPLVNLINKYFP GMFPFKDKFGLFAELNNSSGLFTGFTGVQNI SRIHLVDKWNGLSKVDFWHSDQCNMINGTS GQMWPPFMTPESSLEFYSPEACRSMKLMYKE SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA ESGAMEGETLHTFYTQLVLMPKVMHYAQYV LLALGCVLLLVPVICQIRSQEKCYLFWSSKK GSKDKEAIQAYSESLMTSAPKGSVLQEAKL AHLPDTLLLPPHSPTVPTPKSFOCSOKACFSRS
1033	2383	A	6292	292	167	FCLLLSLVSSSLVSLSLCPPLTQA
1034	2384	A	8597	640	164	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQ PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL VNMGDRTSMVQDPGSQAPTSWISESQVFQTT EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG GAGYVRSSQDLSCDFCNDVLARAKYLKRHG F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI HVYKKNGVGKVGDQILLAIKGQKKKALIVG HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI KTPIPTSLRKREGEYSKVLAIAQNPV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV TQYLQPRSPEECKMFACAKLACTPSLIRAGSR VAYRPISASVLSRPEASRTGEGSTVFNGAQNG VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY AILGFALSEAMGLFCLMVAFLILFAM
1037	2387	A	8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT GMVAHINNSRLKAKGVGQHDNAQNFGNQSF EELRAACLRKGELFEDPLFPAEPSSLGFKDLG PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL PTKNDKLVFVHSTERSEFWSALLEKAYAKLS GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP PQNLLRLLRKAVERSSLMGCSIEVTSDSELES MTDKMLVRGHAYSVTGLQDVHYRGKMETLI RVRNPWGRIEWNGAWSDSAREWEEVASDIQ MQLLHKTEDGEFWMSYQDFLNNFTLEICNL TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV VVCTCLVALMQKNWRHARQQGAQLQTIGFV LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI FTNSREVSSQLRLPPGEYIIPSTFEPHRDADFL LRVFTEKHSESWELDEVNYAEQLQEEKVSED DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR MAIKFKSFKTKGFGLDACRCMINLMBCDGSG KLGLLEFKILWKKLKKWMDIFRECDQDHSGT LNSYEMRLVIEKAGIKLNNKVMQVLVARYA DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT GHICLSLEQVLGEGWEGICRIAPACPSTPPPPS

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ĺ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	neuce	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ŀ	ì	914	ng to first amino scid	acid residue of peptide	Q=Ghttamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	Sequence	/=possible nucleotide deletion, \=possible
		1	1	sequence	•	nucleotide insertion
				· · · · · · · · · · · · · · · · · · ·		SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI
		İ				EAL
1038	2388	Α	8621	3	1494	RSRMARAPLGVLLLLGLLGRGVGKNEELRLY
						HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL
						ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF
						GGIETLRVPSELVWLPEIVLENNIDGQFGVAY
		[DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG
			l i			KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH
1						GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV
	•	İ				LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT
						VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI
1]				VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL
						LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE
		1				ELILKKPRSELVFEGQRHRQGTWTAAFCQSL GAAAPEVRCCVDAVNFVAESTRDOEATGEE
		ĺ	Ì			VSDWVRMGNALDNICFWAALVLFSVGSSLIF
						LGAYFNRVPDLPYAPCIQP
1039	2389	A	8636	1	900	PGRERPGGGGARRRPQHLPALLPSERPDCATL
						QAMENELPVPHTSSSACATSSTSGASSSSGCN
						NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ
						ALHRQPSTAAQYLQQMYAAQQQHLMLQTA
						ALQQQHLSSAQLQSLAAVQQASLVSNRQGST
		ľ	1 1			SGSNVSAQAPAQSSSINLAASPAAAQLLNRA
						QSVNSAAASGIAQQAVLLGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR
			!			PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL
						PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF
						HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ
[]			1 1			RRDSFSGVKDSNNNSDGKAVAKVKCEARSA
						LTKPKNNHNCKKVSNEEKPKVAIGEECRADE
1			1 1			QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG
						NPGSTSAATCTRRHYERLILPYERFIKGEEDKP
						LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP
						KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ
					İ	KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD
						PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA
				ł		PLAPEKDSALVPGASKQPLTSPSALVDSKQES
}						KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP
] '''	~/.	••	55.15		- 1,5-	GIKARITQRALDYGVQAGMKMIEQMLKEKK
				ļ		LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP
						NTSLAFVPGVGIKALTNHGTANISTDWGFESP
	ľ					LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK
						ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE
						NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER
						SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS
					i	TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK
]				į		NSTVETIVSMDFVASTSVGLVILGORLVCSLS
	1		1	ł	ſ	LNRFRLALPESNRSNIEVLRFENILSSILHFGVL
	1					PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF
1						LLISTDLKYETSSKQQPSFHVWEGLNLISRQW
						RGKSAP
1042	2392	A	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR
<u> </u>	1					LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT LDMITSTDHVLEQDFWICFIFYSVKERQI GLKTRAPATPTFQREVLGPAKQDMQRRCPRIGLMTSLLKPIKRRWRDYKRWKSGGFTGESC HHADTLGDRGGLQGDHSELLQWQKRILRTE GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL NLALADLLFALTLPIWAASKVNGWIFGTFLC KVVSLLKEVNFYSGILLACISVDRYLAIVHA TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR RTVYSSNVSPACYEDMGNNTANWRMLLRIL PQSFGFIVPLLIMLFCYGFTLRTLFKAHMGQK HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM RTQVIQETCERRNHIDRALDATEILGILHSCLN PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS RPSFVGSSSGHTSTTL
1045	2395	A	8724	254	3184	FRANLATTVANRRGAQGGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLLQVLILNDNLLSSLPINLFRFVPLTHL DLRGNRLKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQLQPGEEERRSHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFFCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	Α	8736	28	452	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA
1047	2397	A	8741	673	924	ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL
1048	2398	A	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR

NO: of peptide sequence No: of peptide sequence	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Desire D	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
uence main acid	nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
uence 914 maino acid residue of peptide sequence T-Threanine, V-Valiae, V-VII-Typophan,	cotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
mino acid reidue of sequence pritite sequence pupitide sequence publication sequence pupitide sequence pupitide sequence pupitide sequence pupitide sequence pupitide sequence pupitide sequence publication sequence pupitide sequence publication	seq-	uence	l	09/496	correspondi	to last amino	
residue of peptide sequence Poptide sequence	uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
poptide sequence // possible maleotide deletion, l-possible nucleotide insertion VAVPRIGOPPSAARYMPREVPERFECQODIK VLKRGOPPPSCALL GGGAGPPFCTAPGAN PNNAQVTGALLQSEGTAPDSTLGGAAASNY ANSTWGGASSNNGTSPNPHITUNGVIVTOGS DMEEWPCLASKOTESSENTTDNNSASNPGSE KSTLPGSTSNKKGGSGOQSASSGNBCNLGV WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GODRIGPGGGSGSNNSNPHSPNSAWALVOGGTS RKCALETINNISSAQVSTVGGTSLEQQSKME NGGVPVVSGREQAGLENTOGPKNACHTNISL NGSSPNMERNKGMPGMGLGNTSRSTDAPS STORVVVSGREQAGLENTOGPKNACHTNISL NGSSPNMERNKGMPGMGLGNTSRSTDAPS STORVVVSGREQAGLENTOGPKNACHTNISL NGSSPNMERNKGMPGMGLGNTSRSTDAPS STORVVVSGREQAGLENTOGPKNACHTNISL NGSSPNMERNKGMPGMGLGNTSRSTDAPS STORVVVSGREQAGLENTOGPKNACHTNISL NGSSPNMERNKGMPGMGLGNTSRSTDAPS STORVVVSGREQAGLENTOGPKNACHTNISL NGSSPNMERNKGMPGMGLGNTSRSTDAPS STORVDNNGGEWCATTARSPNACHTNISL NGSSPNMERNGGGWGACAPCWGRSS STORVCHOOLD STORT STORVDNNGGEWCAPTCSDELKGGEWGACAPCWGRSS STORVDNNGGEWCAPTCAGNACHTNISL ATSWLADAPS STORVDNNGGEWCAPTCAGNACHTNISL ATSWLADAPS AATOTKNSGGWGAAAACPGGGGGGGACAGAWA AATOTKNSGGWGAAAACPGGGGGGACAGAWA AATOTKNSGGWGAAAACPGGGGGGACAACAACAACAACAACAACAACAACAACAACAA			ľ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
miclootide insertion	1					sequence	
VAVPNGOPPSAARYMPREVPPERFEQQODIK: VLLKRGOPPPSCALLGGGAGPPTCAPGAA PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWIGSGASSNNGTSPNPHIUMVCHVTOGS DMEEWPCLASKUTESSSENTTDNNSASPNGGE KSTLPGSTTSNKGKGSGOQSASSGNECNLGW WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGGGGSFNPNPNSNPAWPLAVQEGTS RKGALETDNNSSAQVSTVGGTSREQQSKME NGVVVSGREQGLGHNTDGFKNGANTNSL NASPNPMENKGMPFGMGLGNTSRSTDAPSQ STGDRKTGSVGSWGAARGPSGTDTVSGGNSS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNSTGGSWPKGPDSNDNN WGGDRK MTSGVSGGEWQPTGSDELKIGEWSGPNQPN STGAWNQKGBFLLENGCHAGPGDSNDN WGGDRK MTSGVSQGEWQPTGSDELKIGEWSGPNQPN STGAWNQKGBFLLENGCAAPCWGRSS SSTGSFVEGGSTGSNIKAGSSGSHNSGRRSV RTHPDCQAVLGTLLSRTDLPPRAJSNTWG QTOKQDTVWDIEEVPRPEGKSDKGTEGWS SSTGSFVEGGSTGSNIKAGSSGSHNSGRRSV RTHPDCQAVLGTLSRTDLPPRAJSNTWG QTOKGDTVWDIEEVRPPGGRSDKGTEGWS SSTGSFVEGGSTGSNIKAGSSGFRASWGELS AATGTKNSGGWGABPSGOMGMKSGWGLLS AATGTKNSGGWGABPSGOMGMKSGWGLLS AATGTKNSGGWGABPSGOMGMKSGWGLS AATGTKNSGGWGABPSGOMGMKSGWGLS SGKNGWGEPPPGVGAWS SGKNGWGEPPPGVGAWS SGKNGWGEPPPGVGAWS SGKNGWGEPPGPGSSRKAMDDDGTSAWG DPGATSSWAMPSKAMCHQQQPPQCPPPPQ PEASGSNGGTPPPNFONNSSSGPQA TFKDEEPSGWEEPSPGSISKAMDDDGTSAWG DPNSTYNCKNNLWAGSQGCAPAGPS WGGABPTWACHTON STANDARG ANDTON STANDARG URSQKKCSLKGGNNGALSAKAGDARGPS WGGABPTCHARGPSTANDARGPS WEEEDGGWNTTGGGGSASSINSASWGGA MLGDPTDMRTCKDSSWAMPLAKGPSNMG LISQTEDNPSSKAMDLSVGSGGPAPERPRIPTP WEEEDGGWNTTGGGGSASSINSASWGGA MLGDPTDMRTCKDSSGRPPNSKOMGTDS GRAGMKCSLKGGNNSSWAMPLAKGPSNMG LISQTEDNPSSKAMDLSVGSGGAPPNSKOMGTDS GRAGMKCSLKGGNNGAMTORS GRAGMATAGATGATAGATGATAGATGATAGATGATAGATGATAGATGAT	1		1		peptide		/=possible nucleotide deletion, \=possible
VILKRGOPPPSCALLIGGGAGPPCTAGCAN PNNAQYTGALIAGSEGTAPDSTLGGAAASNY ANSTWGSGASSNIGTSPRIPEHVDKVIVDGG DMEEWPCLASKUTGSSENTTDNNASAPROSE KSTLPGSTTSNKGKGSQCQSASSGNECALGV WKSDPKAKSVQSSNSTTENNNAGASPROSE KSTLPGSTTSNKGKGSQCQSASSGNECALGV WKSDPKAKSVQSSNSTTENNNAGASVPALVQGGTS RKGALETDNNSSAQVSTVGGTSSEQOSKME NAGVRIVVSGREQAGEINTDGFKAGNTINSL NLSSPPMERMCKOMPFGMGLENTISTDAPSQ STODRICTGSVGSWGAARGPSGTDTVSGQSSK WNNDRSTGGSWPGAARGPSGTDTVSGQSNS GPNGENGKEREDSVKKGASVQKSTGSKJDS) WNNDRSTGGSWGAARGPSGTDTVSGQSNS GPNGENGKEREDSVKKGASVQKSTGSKJDS WNNDRSTGGSWGAARGPSGTDTVSGQSNS GWNGNGKERELBROGTGSNEKGGSWGAARGPSGTDTVSGQSNS GWNGNGKERELBROGTGSNEKGGSWGAARGPSGTDTVSGQSNS GWNGNGKERELBROGTGSNEKGGSNSGNGNS WNNDRSTGGSWGFGOSDELRIGEWGGFNGWGNS STGAVUNOKGEPILLENGCAAAPCWGRSS STGSEVEGGTGSNEKAGSSDSHNSGGRSY RTHPDCQAVIQTILLSRGTDLDPRVSINTGWG QTQKKQTTVDIEEVPRPEGKSDKGTEGWSS STGSEVEGGTGSNEKAGSSDSHNSGGRSY RTHPDCQAVIQTILLSRGTDLDPRVSINTGWG QTQKKQTTVDIEEVPRPEGKSDKGTEGWSS SIGKNGWGEDVGTKNSNWSSASKWGLS AATGTKNSGGWGDRFDROGWGGRGPNGGWS SIGKNGWGEDVGTKNSNWSSASKWGS GGGREITOWRGGNASLASKGGWEGELS AATGTKNSGGWGDRFDROGWGGRGPNGGWS SIGKNGWGEDVGTKNSNWSSASKWGS BGGREITOWRGGNASLASKGGWEDCKSS PAWNETGRGPNSWNRQHQQQPPQOPPPP PEASGSWGGPPPPFONVRYSNSNWSSGCPAREPNLPTP MTSKSASDKSKMQDGWGESDCYTCARHPS WEEEDGGVWTTTOSQASSSINKSGCPAREPNLFTP MTSKSASDSKSMQDGWGESDCYTCARHPS WEEEDGGVWTTTOSQASSSINKSYGPOKA LLSGTEDNPSKMDLSSGGFAPPNSKDMGTTDS GYFEKGGSHGLFONTGAGSSASTHVORA MILGDFNDIMRKDRSGGRAPPNSKDMGTTDS GYFEKGGSHGLFONTGAGSGSHTVOPKN MILGDFNDIMRKDRSGGRAPPNSKDMGTTDS GYFEKGGSHGLFONTGAGSGASHTVOPKN MILGDFNDIMRKDRSGGRAPPNSKDMGTTDS GYFEKGGSHGLFONTGAGSGASHTVOPKN MILGDFNDIMRKDRSGGRAPPNSKDMGTTDS GYFEKGGSHGLFONTGAGSGSHTVVHSTSAK POPKSTRAGAGPTPTDTDHOLLDNTGS NSSLNTTLERSGGAWPYSASDNSTTNVHSTSAK POPKSTWSPDPIGHNTHLSNCMWCHNISS RNTTPLRRPPGLTNPRNSSSWSTAPRSVRG WGTODSRLAAGSTWSPDGORPMSSSTRVVHSTLAAR PDYKSTWSPDPIGHNTHILSNCMWCHNISS RNTTPLRRPPGLTNPRNSSSWGFTPVFRAFSSAVLIT LITQJGGSTLRTICMQFRTTSSSLWQFTYTED PHRMGSPRILLAGCFTVDTSSSLWCTTLLLISTFFLAPWDS NGCREGGRAPHYNTSSLWGFTYNFSLWPTYFE PHRMGSPRILLAGCFTVDSSSACTLDSSFPTLAPWDS OGSSGADLAGASLWGPRYNSSLWCTTLLLILLSTLT LETILLLAGASLVWWA					sequence		nucleotide insertion
PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWIGSGASSNNGTSPIPRIHUNDKVIVTOGS DIMEWPCLASKDTESSENTTDNNSASPROGE KSTLPGSTTSNKGKGSGOQSASSENECNLGV WKSDPKAKSVQSSNSTTEINNIGLGWRRIVS GODRIGGGGSGESNPINSBVSAVPALVOGGTS RKGALETDNNSSAQVSTVGQTSLEQQSKME NAGVAFVVSGRQAQIHNTDGPRAGNTINSL LSSPPPMBNKGMPFGMGLOTISSTDAPSQ STGDRKTGSVOSWGAARGPSGTDTVSQGNS GONGRINGSKEDSWKGASVCKSTGSKNDS WDINNIRSTGGSWNFGPQDSNDNKWGEGNIK MTSGVSGEWKQPTGSDELLIGGEWGSPOPN SSTGAWDNOKGHELLENQGAQAPCWGRSS STGSEVEGOSTGSNIKAGSDGSNFSGNSOS WDINNIRSTGGSWNFGPQDSNDNKWGEGNIK MTSGVSGEWKQPTGSDELLIGGEWGPOPN SSTGAWDNOKGHELLENQGAQAPCWGRSS STGEVEGOSTGSNIKAGSDGBNSWGRGUS ATTEWCPPTGCAVALOTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPPFGKSDKGTGGWS AATQTKNSGGWGAPSPGNOKWSGWGELS ATTEWCPPTATSSWNESPSKDQGWGGRQPNQGWS GKNGWGEWDQTKNNNWSSASSPVSWG GGGQNEIGTWGNGGRQPNQGWS SKIKGWGEWDQTKNSNWSSASSPKYSWG GGGQNEIGTWGNGGGRQPNQGWS SKIKGWGEWDQTKNNWSSNSWSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSGCPQP PEXSGSWGGPPPPPFGNVRFNSSWSSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSGOPQP PEXSGSWGGRPPPPPFGNVRFNSSWSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSGOPQP PEXSGSWGGPPPPPFGNVRFNSSWSGOPQP PEXSGRAFGAGTGSPFFQCWFP PHDSMCMARDASGRPPPNSSWMGGGGRQFNGMHSSPFFN WEEEEDGGVWNTTGSQGSASSHNSASWOQG GKQMKCSIKGGNIDSWMPPLAKQFSNMG LLSQTGDMFSSKMDLSOPQFGAKGSPFPNCFF PHLDNMYNALNOLPPLOTKGPPFQYGGG SGGGMDAMAGNGAGCTSSRFCQWFSMME GLPSVATQEAMMHKNGAIVAPGKTRGGSPY NGPDIPGDTLGGHTGAGDSWLPAKGFPNNHS SSSLRTAQVPPGISPQVSASMSTAVRGQGSQCQL LLQQQQQQLLQNQRKINGASPPTNK GSKSSNASWPPEPQFQFWKGQNIDPESDP VYTGSVLGATASTPVTDTHQLLADNTTGA SNSLNTTLPSFGAWPYSASDNSTTNVHSTSAK PDVKSTWSPDPIGNFTHLSNKMWNHSS RNTTPLPRPPGLTAPHCADSWLPATAGCWSSA GGSSGADLAGASIWGPTNASSSWQVFTVED PHRMGSPALLFUCHJELBWTLTLLLSTFT LPPSSLFPFLQVDSSSSACTLDSFFFTLAPWDS LETQQDGSDPVQPALALFTGASSVALUTH LTYSTKGGAAKATALHACVLGGTTLAEF ATDDEVSR-LAQAQPTPAATPSAPAACWG LETQQDGGDLGALTWTLLLLSTTT LPPSSLFPFLQVOSSSACTLDSFFFFLAPWDS POCGFGROPQCLSLTVYTHILLLS TAMYGAHAPLLALCHVDGGVPF			-				VAVPNGQPPSAARYMPREVPPRFRCQQDHK
ANSTWIGGRASSINGTSPIPHIUMDKYVTORG DIMERWICHASKIPTISSSENTITINNIASNPIGGE KSTLPGSTTSINKGKGSQCQSASSGNECHLQIV WKSDPKAKSVQSSNSTTENNIGLGINWRNVS GQDRIGPGSGFSNFMPNISSAWPALVQEGTOR RKGALETIDNISSASQVSTVGGTSEQOSKIME NAGVNPVVSGREQAQHINTDGFKNGNTINSL MLSSPIPMENKGMPFGMIGLGNTSTADASQ STDDRKTGSVGSWGAARGFSGTDTVSGQSSS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVVSKTGSKDNS GNANDNIGKERESWKGASVASTADAS GNANDNIGKERSWKGASVASTADAS STGSEVEGGSTGNHEKAGSSDSHNSGRKSV RTHPDCQAVLQHLSRTDLIPRQAPCWGNS SSTGSEVEGGSTGNHEKAGSSDSHNSGRKSV RTHPDCQAVLQHLSRTDLIPRGWS GTAKKNSGWGDAFVSKNMKGMSGWEGS AATGEKRANDNIKGMGGAPTUSNTGWG GTAKKNNGGWGDAFSSRNMKGMSGWGELS AATGEKRANGGWGDAFSSRNMKGMGGG PDEKTFSSWWEIPSDQGWGGGRQPNGGWS SGKNGWGGEVDQTKNNSWESSASKYSGWQ GROORIGTWGNGGNASLASKGWGDCKRS PAWNETGRQPNSWNKGHQQQPPQQPPPQPPP PRASGSWGGPPPPGRAFFRSNSWSGPQPA TFKDEEFSGWEEPSSQSISRKMDIDDGTSAW GDPNSYNTYNNVLWKNNGQGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPYTGARFPS WEEEEDGGVWNTTYSQGAASSINSSWSGPQPA TFKDEEFSGWEEPSSQSISRKMDIDDGTSAW LIQTENPSSKUMDLIVSGSGKKFDVDKRA MILGPENDBIRKDRJSGSKKFDVDKRA MILGPENDBIRKDRJSGSKKFDVDKRA MILGPENDBIRKDRJSGSKKFDVDKRA MILGPENDBIRKDRSGFFPPNSKDMGTTDS GPYFFKGGSBGLGFSSTAVSGGGKGGKGKKFDVDKRA MILGPENDSKGMGLSVGSSSKKFDVDKRA MILGPENDBIRKDRSGFFPNSKDMGTTDS GPYFFKGGSBGLGFSSTAVSGGKGKGGTSSFSFVVAR LIQQQQQOLLUNGNISISQAVRQQGEQLAL RKVGKGGKGGGGGMGGMGMSSSFSFVVAR RKGGRGGSGGGMGMGWGMGMSSSFSFVVAR RKGGRGGARAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA	1		İ	}			VLLKRGQPPPPSCMLLGGGAGPPPCTAPGAN
DMEEWPCLASKDTESSESNTIDNNSASNPGE KSTLPGSTTSNKGKGSQQCASSGGECRLGV WKSDPKAKSVQSNSTITENNGLGNWRNY GQDRIGPGGFGFNPNSNPSAWPALVQEGTS RKGALETDNSNSSAQVSTVGGTSREQOSKEM NAGVDFVVSGEGAQHIPTOGPKGATINSL NAGVPVVSGEGAQHIPTOGPKGATINSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ STODIKTGSVGSWGAARGPSGTDTVSGGSSNS GNNGNGKEREDSWKGASVQKSKTGSKNSNS WDNNNRSTGGSWNFGPQDSNDKKWGEGNN MTSGSVGGWKQFFGSDELKIGEWSGFNOPN SSTGAWDNQKGHPLLENCGNAQAPCWGRS GTOKKQTVVWDEEVPFECKSKNSKGRSSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGW QTOKKQTVVWDEEVPFECKSKNGKGSWS STGGWDGWQTQTULSRTDLDPRVLSNTGWG QTOKKQTVVWDEEVPFECKSKNGKGSWS STGAWDNQKGHPLLENCGNAQAPCWGRS STGAWDNQKGHPLLENCGNAQAPCWGRS STGAWDNQKGHPLLENCGNAQAPCWGRS STGAWDNQKGHPLLENCGNAQAPCWGRS AATQTKNSGGWODAPSGNQMKSGWGELS AATGTKNSGGWODAPSGNQMKSGWGELS SGKNGWGEVDGTNSNNWSSASKAFVSGWG GGONEIGTWGNGGNDSVNGHAGGOR SGKOWGEVDGTNSNNWSSASKAFVSGWG GGONEIGTWGNGGNNDTNNNNSSNWGGGR PDEKTTSSWNENPSKDQGWGGGRQPNQGW SGKNGWGEVDGTNNSNWSSASKAFVSGWG GGONEIGTWGNGGNNSLASKGGWEDCKAF PAWNFEGROPSWNKGHQQQPQCPPOPPP PEAGGSWGGPPPPPGPGNTRSNSSWSSGPOP TRYLDEEPSGWEEPSSQSISKMDDDGTSAWG DPNSYNYXNVALWDKNSQGGAPREPPLIPT MTSKSASDSKJMGDGWGGSGPCPTGARIPS WEEEDGGWNTTTGSGGSASSINSASWGGG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKCSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKKGMKGSLKGGNDSWMDPLAAQSWONG GKGMCGGRONGSMAGAGTGERKGWTSMMG GGRAAGATALMCCTGATTLAEF AMDAGATATAGAAAAACAAAAAAAAAAAAAAAAAAAAAAAAA	1	[PNNAQVTGALLQSESGTAPDSTLGGAAASNY
KSTLPGSTTSNKGKGSQCSASSGMECHLGY WKSDPKAKSVQSSNSTTENNIGLGNWRNVS GQDRIGPGSGFSNFMPNISPSAWPALVQEGIS RKGALETDNSNSSAQVSTVGGTSREQGSKME RAGWPVVSGREQAQHINTDGPKNGTINSL NLSSPNPMENKGMPFGMGLGNTSRTDAPSQ STODKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNSTGGSWNFGPQDSNDNKWGEGINK MTSGVSQGEWKQPTGSDELKIGEWSGPNGPN STGAWDNQKGPHLENQGNAQAPCWGRSS STGSEVEGGSTGSNHKAGSSDSIHNSGRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWESS AATQTKNSGGWGDAPSOSNQMSGWGLS AATEWKDPKNTGGWDTYKNNISSNWGGG PDEKTPSSWMEDSPSDQGWGGGRQPNQGWS SGKNGWGEEVDQTRNSNWESSASKPVSGWG EGGWEIGTWRGGNASLASKGWEDCKRS PAWNETGRQPNSWNKQHQQQPPQOPPPD PASGSWGGPPPPPGNVRPSNSSWSGPQPA TPKDEEPSGWEEPSPGSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPEEPHLTP MTSKSASDSKSMGDGWGGGGNSGMAGMGGGGGGGGGGGGGGGGGGGGGGGGGGGGG							ANSTWGSGASSNNGTSPNPIHIWDKVIVDGS
WKSDPKAKSVQSNSTITENNIGLGNWRNYS GQDRIGPGGGFGNPNSNPSNSPAWPALVQEGTS RKGALETINSNSSAQVSTVGGTSREQQSKME NAGVNVVSGREQAQHINTOGPKGGNTNSL NASSPNPMENKGMPFGMGLGNTSSTDASSQ STDDRKTGSVGSWGAARGPGGTDTVSGQSNS GDNGNNGKEREDSVKGASVQKSTGSKNDS WDNNNSTGGSWGAKGPGDSKDJKKWGGGNN MTSGVSQGWKQPTGDSKDJKKWGGGNN MTSGVSQGWKQPTGDSKDJKKWGGGNN MTSGVSQGWKQPTGDSKDJKKWGGGNN MTSGVSQGWKQPTGDSKDJKKWGGGNN STGAWDNQKGPPLLENCGNAQAPCWGSS STGDRKTGSVGSTSHKAGSSDSHNSGGRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTOWG QTQIKKQDTVVDDESVPREGKSDSHNSGGRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTOWG AATQTKNSGGWGDAPSGSNQMKSGWGELS AATGTKNSGGWGDAPSGSNQMKSGWGELS AATGTKNSGGWGDAPSGSNQMKSGWGELS SGRNGWGEEVDGTKNSNWESSASKPVSGWG GGQNEIGTWGNGGNADLASKGGWBCKGCS SGRNGWGEEVDGTKNSNWESSASKPVSGWG EGQQNEIGTWGNGGNASLASKGGWBCKCK SGRNGWGEEVDGTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWBCKCK PPEKTPSSWKENPSKDQGWGGGGPNGPNQPN PRASGSWGGPPPPPGFWRTSNSSWSSGPOP PRASGSWGGPPSPPPGRWFRSNSSWSSGPOP PRASGSWGGPPSPPFGNTSRNSWSSGPOP PRASGSWGGPPSPPFGNTSRNSWSSGPOP PRASGSWGGPPSPPFGNTSRNSWSSGPOP PRASGSWGPPSSSMSTANGWGGG GKKQMKCSLKGGNDSWMDN-LACPSNMG LLSQTEDNFSSKNDLSVGSLSDKKFDVDKRA MNLGDPNDIMRKDRSGGPARPSNLAGFSNMG LLSQTEDNFSSKNDLSVGSLSDKKFDVDKRA MNLGDPNDIMRKDRSGGFRPSSNALTH GGSTSANGWGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG		1					DMEEWPCIASKDTESSSENTTDNNSASNPGSE
GORDIGPGSGESIMPINISMSAWPALVGGTS RKGALETDNISSAVOSTOREQOQSME NAGVIPVVSGREQAQDENTIDGPKNGNITISL NLSSPIMPEME KOMPFGMGLOTISSTDAPSQ STODRXTGSVOSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNRTGGSWHGPGVDSSDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNGPN SSTGAWDNQKGPHLENGGNAQAPCWGRSS SSTGSEVEGGSTGSNIKKAGSSDSINNSGRRSY RPTHPDCQAVICTLISSTDAPSQ QTQIKQDTVWDIEEVPRPEGKSDKGTGWES AATQTKNSGGWGDAPSGSNMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNAGGR PDEKTPSSWNENPSKDQGWGGRQPNQGWS SGKNGWGEVDOYKNNNWSSSASKPVSGWG EGQONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGW GEGCONEIGTWGNGGAASASKPVSGWG GEGCONEIGTWGNGGAASASKPVSGWG GEGCONEIGTWGNGGAASASKPVSGWGCACAA ANGLOPHDIMRKDRSGFRPPRSNSWSGOPPPP PEASAGSWGGGPPPPPPGNAVPFSNSWSGGPAAPEPPLITP MTSKSASDSKAMQDWGGGAASASKPVSGWGCACAA GEGCONEIGTWGNGGAASASKPVSGGGAAPEPPHLITP MTSKSASDSKAMQDWGGGAASSHNSASWGQ GKKQMKCSLKGGNNDSWMNLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKRDVDKRA MNLGDPNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSILAQVPQPJSPQVSKAMIKOFPNSGLSP GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSILAQVPQPJSPQVSKAMIKOFPNSGLSP GPYFEKGGSHGLANNIKAGAACAACAALMCPGPDYGGSG SGGMDYGMVGGKAAGTESRFKQWTSMME GLPSVATQAAMAKAACAACAALMCPGPDYGSGA WGTQASALASASTWGGGGAACACAACAALMCTATSAKAACAACAALMCTATSAKAACAACAALMCTATATSAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA							KSTLPGSTTSNKGKGSQCQSASSGNECNLGV
RKGALETDNSINSSAQVSTVQGTSREQQSKME NAGWFVVSGREQQB,NTDGFKNGMTNSL NLSSPIPMENKGMPFGMGLGHTSRSTDJPSQ STODRXTGSVGSWGAARUFGSTDTVSQGSNS GNNGNMGKEREDSWKGASVQKSTGSKODS WDNNNRSTGGSWGAARUFGSDELKIGEWSGPNQPN STGAWDNQKGHPLLENGGNAQAPCWGRS SSTGSVEGGSTGSNIKKASSDSINNSGRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQGKQDTVWDHEVPRPEGKSDKGTEGWES AATOTKNSGGWGDAPSOSNQMKSGWGLS ASTEWEDPKNTGGWNDYKNISSWGGLS ASTEWEDPKNTGGWNDYKNISSWGGLS ASTEWEDPKNTGGWNDYKNISSWGGLS ASTEWEDPKNTGWNDYKNINSSWGGGR PDEKTFSSWNENPSKDQGWGGRQPNQGWS SGKAGWGEEUDQTKNSNWSSASRPVSGWG EGGQNEIGTWGNGNASLASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQPPQQPPPP PRASGSWGGPPPPFFONVRPSNSWSSGPQPP TFKDEEPSGWEEPSPQSISKKMDIDDGTSAWG DPNSYNYKNYN,WDKNSQGPAPREPNLETPP MTSKSASDSKSMQDGWGESDQPVTGARHPS WEEEEDGGWNTTTOSQGASSHNSASWGG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDMPSKMDLSGSHDAFNSWSSHNAGHQ CLSQTEDMPSKMDLSGFPPPNSKMDGTTDS GPYYFEKGGSHGLFGNTSAGSLHTVOPLN SPPSLRAQVPPQFISPQVSASMLKQFPNSGLS GLINVOPPCLSPQQLANLSQLPQFQCIACQL LLQQQQQQLDNRKKISPAGLATPVOFL ARMYSALQQQQQQQQCPGMKHSPHPVGFK PHLDNNYPHALNVGLPDLQTKGPFGYGSGF SGGMDYGMVGGKEAGTESRKQWTSMME GLPSVATQEAMHGNGATVAGGKTGGSFY NQFDIIFDDTLGGHTGPAGDSWLPAKSSPTINK IGSKSSNASWPPEPCPQFVPWKGIONIDPPSDP YYTPGSVLGGTATSPIVDTDHOLLRDNTTGS NSSLNTSLPSPGAWPYSADNSFTNVHSTSAK FPDYSSTWSPPIGRAPPTINSNKWKNNISS RNTTFLPRPFGLTNFKRSSPWSSTAPRSVRG WGTODSRLASASTWSDGGSVAPSYWLVLIN LITPQDGSTLRTICMQHGPLLTFHLNLTQGTA LIKYSTQCAAKAQTALHMCHGTLAGTHTLAFT ATDDEVSRFLAQAQPTPAATTSAPAAGWGS LETGQNOSDPVQPALNLFGSTLQGWSSSA GGSSGALDAGASLWGPPTYSSSL WGYPTVED PHRMGSPAPLLPGOLLOGGNTTLLSFTF LPPSSLPFPLQVDYSSACLLDSFTPFLLAFWTT LPPSSLPFPLQVDFQPALLFGTGLQWSSSA GGSSGALDAGASLWGPPTYSSSL WGYPTVED PHRMGSPAPLLPGOLLOGGSCAULTILESTTF LPPSSLPFPLQVDTSSSACLLDSFTPFLLAFWTT LPPSSLPFPLQVDTPALLSRTTLLAFWTMLLLSTTF LPPSSLPFPLQVDFLALARWTTMLLLSTTL LPPSSLPFPLQVDFLALARWTTMLLLSTTL LPPSSLPFPLQVDFGARAGTLLTPRATTSAPAAGWGS LETGQNQSDPVQPALLAFCTGLAGWSSACLLT EKKLLLCARSLLVWGWAPWPGPPPWRQAA		1					WKSDPKAKSVQSSNSTTENNNGLGNWRNVS
NAGVIFVVSGREQAQIENTIDGIFKNGNTINS. NLSSPNPERIKGMPFGMGLTRSSTDAPSQ STODRKTGSVOSWGAARGPSCTDTVSGQSNS GNNGNMGKEREDSWKGASVQKTGSKODSNS WDNNNRSTGGSWHGPQDSNDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGPHLENQGNAQAPCWGRSS SSTGSEVEGQSTGSNIKAGSSDSINSGRRSY RPHPPDCQAVLOTILSRTDHVSINTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATOTIKNSGGWGBAPSGONKSGWGEIS ASTEWKDPKNTGGWNDYKNNINSSNWGGR PDEKTPSSWIENPSKDQOWGGRQPNQGWS SGKNGWGEVDQTKNNNWESSASKPVSGWG EGQQNEIGTWGNGGNASLASKGOWEDCKRS PAWNETGRGPNSWNKCHQQOPPQCPPPPQ PEASGSWGOPPPPPGNVRPSNSSSGPPA TPKDEPPSGWEEPPSQSISKKDDGTSAWG DPNSYNYKNVILWDKNSQGGPAPEPPHPP MTSKSASDSKSMQDGWGEGPPPPFONVPSNSWSGPPA PAWNETGRGPNSWNKCHQQOPPQCPPPPQ PEASGSWGOPPPPPGNVRPSNSWSGPPA PRYSKASDSKSMQDGWGEGPAPEPPHPP MTSKSASDSKSMQDGWGEGFSDKFRDVDKRA MNLGDPNDIMRKDRSGGFRPPNSCMMGTTDS GPYPEKGGSHGLFONTSTOSRGLHTPVQPLN SSPSLRAQVPPQTSRPVSASMLKQPTPNSGLSP GLPNVOPQLSPQOLANS GLPNTPNSGLSP GLPNVOPQLSPQOLANS GLPNTPNSGLSP GLPNVOPQLSPQOLANS GLPNTPNSGLSP GLPNVOPQLSPQOLANS GLPNTPNSGLSP GLPNVOPQLSPQOLANS GLPNTPNSGLSP GLPNVOPQLSPQOLANS GLPNTPNSGLSP GLRNVAPQLSPQOLANS GLPNTPNSGLSP GLRNVAPQLSPQOLANS GLPNTPNSGLSP GLRSVATGEAMMENGALVAPGKTRGGSPY NQPDIIEDDTLGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPEPPPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTLSPSGAWPYSASDLVAPGKTRGGSPY NQPDIIEDDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPEPPPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSSNASWPEPPPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTLSPSGAWPYSASDLVAPGKTRGGSPY NQPDIIEDDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPEPPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSSNASWPEPPQPQPPPWRGQMPPWSSSLWGYTYED PHRIMGSPAPLLPGLLTPRKTSPPSSTAPRSVRG WGTQDSRLASASTWSDQGSVRPSYWLVLHN LTPQDGSTLRTICAMPGTATGGTAPWSSSLAGGWGS GNSGADLAGASLWGPPTYSSSLWGYTYED PHRIMGSPAPLLPGOLLOGGSGGGGGGSGGGWSGAGUTLASTFTLSPTF LPPSSLPFFLQVIDSSSSACTLDSFTFRLAFWBS PQDCGFKDHQPLTLQALTVELARWTTMLLISTF LPPSSLPFFLQVIDSSSSACTLDSFTPALTSAPAA]					GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS
NLSSPPMENKGMPFGMGLGHTSRSTDAPSQ STODRKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKADS WDNNNRSTGGSWNFGPQDSDNMKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRS SSTGSEVEGGSTGSNIKAGSSDSINNSGRSY RPTHPDCQAYLQTLLSRTDLDPRVLSNTGWG QTQICKQDTVWDIEEVRPERSADKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS ASTEWADPKNTGGWNDYKNISSWGGGR PDEKTPSSWNENPSKDQGWGGRQPNQGWS GRQCNGEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGASLASKPVSGWG EGGONEIGTWGNGGGAPREPNLETP MTSKASADSKSMQDGWGESDQPVTGARHPS WEEEEDGGWWNTTOSQGASSHNSASWGGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSGTEDNPSSKMLDGGGAPREPNLETP MTSKASADSKSMQDGWGESDQPVTGARHPS WEEEEDGGWWNTTTOSQGASSHNSASWGGQ GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSGTEDNPSSKMLDLSGAFPNSKLMGTTDS GPYFFEKGGSHGLFGNTSTAGSLFWYDDKACH LLSQTEDNPSSKMLDLSQLFPOPCALCQL LLQQQQQQLLLONQKISGAVRQQEQQLACQL RMVSALQQQQQQQLALNISQLPQPFQLACQL LLQQQQQQLLONQKISGAVRQQEQQLACQL RMVSALQQQQQQQDQFGMKHSPSHPYGFK PHLDNNVPNALNVGLPDLOTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEAMMKNGALVSQLFQPFQTSASK GISPSVLGGTATSPIVDTDHQLLRDNTTGS NSSLATSLPSPPEQPGAWPYSASDRYWSTARKSVG WGTQDSRLASASTWSDGGSWPSYWLVLIN LTPQDGSTLRTICMQHOPLLTPHLNLTQGTA LIKYSTKQEAAKQTALHMCLGNTTLLAFT LTPQDSTRLAQAQPTPAATSAPAAGWGS LETGQNQSDYPALNLTFGTLQGWSSA GSSGALDAGASLWGPPTYSSSLWGVPTVED PHRMGSPAPLLPOQLLAGGSTLQGWSSA GSSGALDAGASLWGPPTYSSSLWGVPTVED PHRMGSPAPLLPOQLLAGGSTLGQWSSA GSSGALDAGASLWGPPTYSSSLWGVPTVED PHRMGSPAPLLPOQLLAGGSTLARWTIMLLS TAMYGAHAPLLALCHYDGRYPFRSSACLLDSFFPFLAPWDS PQDCGFKDHQPLTALATCHLARWTIMLLS TAMYGAHAPLLALCHYDGRYPFRPSSALLT ETKLLLCASTLLVWGAWPGPPPWRQAA							RKGALETDNSNSSAQVSTVGQTSREQQSKME
STGDRKTGSVGSWAAARGPSGTDTVSGQSNS GNNCHNIGKEREDSWAGAXQKSTGSKNDS WDNNNRSTGGSWNFCPQDSNDNKWGEGNK, MTSGVSQGWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGPHLEHOQRAQAAPCWGRSS SSTGSVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVIQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATOTKNSGGWGDAPSQMKSGWGELS ASTEWKDPKNTGGWDDYSNNNSSWWGGGR PDEKTPSSWHENPSSDQGWGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPV30WG EOQNEIGTWONGGNASLASKGGWEDCKRS, PAWNETGRGPNSWNKQHQQQPPQOPPPPQ PEASGSWGGPPPPPGMVRPSNSSWGGPPPPPGPPGPPGPPPQ PEASGSWGGPPPPPGMVRPSNSSWGGPPPPPGPPPGPPPGPPPPGPPPPGPPPPGPPP	1			ĺ			NAGVNFVVSGREQAQIHNTDGPKNGNTNSL
GNIGNIGEREDSWEGASVQKSTGSKNDS WYDNINERSIGSWNFGPODSHDNIKWGEGNE WTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGRAQAFUCRESS SSTGSEVEGSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVVDIEEVPPFGKSDKGTEGWES AATQTKNISGGWGDAPSQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGG PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTIKNSIWESSASKPVSGWG EGGQNEIGTWGGGNASLASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQQPQPPPPPPP PEASGSWGGPPPPPGNTVPRSNSSWSGGPQPA TPKDEEPSGWEPSPQSISRKMDIDGTSAWG DPNSYNKYNVLLWDENSAWSGGPAPEEPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWQQG GKKQMKCSLKGGNNDSWMPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDDGTSAWG MILGSTEDNPSSKMDLSVGSLSDKKFDDKTA MILGSTEDNPSSKMDLSVGSLSDKKFDDKTA MILGSTEDNPSSKMDLSVGSLSDKKFDDKTA MILGSTEDNPSSKMLSVGSLSDKKFDDKTA AMILGDFDIDMRKDRSGFPPPNSKDMGTTDS GPYFEKGGSIGLFGNSTAQSRGLHFVQFLN SSPSLRAQVPPOFISPQVASMLKQFPNSGLSP GLFNVGPQLSPQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ	1	}					NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ
WDNNNRSTGGSWNFGPQDSNIDNKGEGNK MTTGGVSQGEWKQPTGSDELKIGGWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRSS SSTGSEVEGGSTGSNHKAGSSDSINGGRSY RPTHPDCQAVIQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNISGGWGDAPSQMKASGWGELS AASTEWKDPKNTGGWDDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS, PAWNETGRQPNSWNKQHQQPPQQPPDPPP PEASGSWGGPPPPPGNVRSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISAMDDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSIKSASDSKSMQDGWGESDQPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQ GKKQWKCSLKGGNNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGSASSHNSASWGQ GKKQWKCSLKGGNNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGSASSHNSASWGQ GKKQWKCSLKGGNNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGSASSHLKGPNSGLSP WEEEEDGQUNTTGSQGASSHNSASWGQG GKKQWKCSLKGGNNDSWMNPLAKQPSNMG GLBYWFQGSGGSGSGSGSGSGSGSGGSASHNSASWGQ GKQWKGGLKGQWNTTGSQGASSHNSASWGQ GKQWKCSLKGGNNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGTAGAHPS WEEEEDGQULAQNGPGSSASHLKGPNSGLSP GRYFEKGGSHGFGNSTAQSRGGHTPVQFLN SSPSLRAQVPPQFISSQVSASMLKQPFNSGLSP GLPTWGPQLSPQQIAM_SQLPPPPQFLACQL LLQQQQQQQQRQFGMKHISPSHPYGPK PH.DNMVPNALNVGLEJDTGKPPGVSGG SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGANVAPCKTRGGSPY NGFDIPRODITLGGHTTGPAGNUPASSHPTNK IGSKSSNASWPPEPQPGVPWKGGNDNEDSDP YVTFGSVLGGTASSPLTDHOLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNWISTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLTNRFPSSPWSTARSVRG WGTQDSRLASASTWSDGGSVRSYWJVLIHN LITPQIBGSTLRTICMGHGPLTFHILNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTAATPSAPAAGWG UCTQDSRLASASTWSDGGSVSTWYVLIHN LITPQIBGSTLRTICMGHGPLTFHILNTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTAATPSAPAAGWG UCTQDSRLASASTWSDGGSSSI PPRMGSSPAPLIPGDLIGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYJDSSVAUCDSPSPALVEDP PPRMGSPAPLIPGDLIGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYJDSSAVHLLSPTF LPPSSLPPFLIDDSSSSACLDSFPPRAWDS PQDCGFKDHQPLTTQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRYPFRESSAVLIT ELTKLLLCAFSLLJWGWAWQQPPPWKQAA	1	İ		1			STGDRKTGSVGSWGAARGPSGTDTVSGQSNS
MTSGVSQGEWKQPTGSDELKIGEWSQPNQPR SSTGAWDMQKGHPILENGAQAPCWGRS SSTGAWDMQKGHPILENGAQAPCWGRS SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQTKQDTVVDIEEVPREGKSDKGTEGWES AATQTKNISGGWGDASQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWHENPSKDQGWGGGRQPNQGWS SGKNGWGGEVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQQPQPPPPP PEAASGSWGGPPPPPGNTYPNSSNWSGGPAPEPNLPTP MTSKSASDSKSMQDGWGESDQPVYGARHPS WEEEDGQCWNTTTGSOGGPAPEPNLPTP MTSKSASDSKSMQDGWGESDQPVYGARHPS WEEEDGGVWNTTGSOGGAPAFEPNLPTP MTSKSASDSKSMQDGWGESDQPVYGARHPS WEEEDGGVWNTTGSOGGAPAFENLPTP MTSKSASDSKSMQDGWGESDQPVYGARHPS WEEEDGGVWNTTGSOGGAPAFEPNLPTP MTSKSASDSKSMQDGWGESDQPVTGARHPS WEEEDGGVWNTTGSOGGAPAFEPNLPTP MTSKSASDSKSMQDGWGESDQPVTGARHPS WEEEDGGVWNTTGSOGGAPAFEPNLPTP MTSKSASDKSMQDGWGESDQPVTGARHPS WEEEDGGVWNTTGSOGGAPAFEPNLPTP MTSKSASDKSMDLSVGSLSDKKFDVDKRA MNLGPFNDINRKDRSGFPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSKGLHTPVQFLN SSPSLRAQVPPPFGSSASMLKQFPNSGLSP WEEEDGGVWNTTGSOGAPAFPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSKGLHTPVQFLN SSPSLRAQVPPPFGSSASMLKQFPNSGLSP GLFNVGPQQQQQQQQQQQQQQQQQQQAQ GRACHKBARACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1						GNNGNNGKEREDSWKGASVQKSTGSKNDS
SSTGAWDNOKGIPLLENQCNAQAPCWGRSS SSTGSEVEGGSTGSINEAGSSBAHNSGRRSY RPTHPDCQAVLQTILISRTDLDPRVLSNTGWG QTQIKQDITVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPQSSNQMKSGWGELS AATQTKNSGGWGDAPQSSNQMKSGWGELS ASTEWKDPKNTTGWNDYKNNNSSNWGGGR PDEKTPSSWEEPPSGQWGGRGRQPPQGWS SGKNGWGEEVDQTKNSNWESSASKPVSQWG EOGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGGPNSWNKOHQQQQPQQPPPPQ PEASGSWGGPPPPPGNVRPSNSWSSGPQPA TPKDEEPSGWEEPPSGSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGWNTTTOSQGSASSHNSASWGQG GKKQMKCSLKGGNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFPVDKRA MNLGDPNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHG_FGNSGRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQGQLAMLSQLPQIPQFGLACQL LLQQQQQQQQLQNQRGMKHSQAVRQQQQQLA RMVSALQQQQQQQQQQQGPGMKHSPSHPVGPK PHDDMWPNALNVGLAGSGFRPYNSGLSP GLFSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIEGDTLGGHTGPAGDSWLPAKSPFTNK IGSKSSNASWPPPGPGPVPKKGQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK PPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLTNPFRSPWSSTAFRSVRG WGTQDSRLASASTWSDGGSVRSYVLVLHN LTQUIDGSTLRTICMQHGPLLTFHILNLTQGTA LIKYSTKQEAAKAQTALHMCVLGNTTILAFF ATDDEVSRFLAQAQGPTAATPSAPAAGWQS LETQQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSTRALPGGDLJGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLLDSFFPTLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYQAHAPLLACHSVDGKVPPRPSSSXLWGVPTVED PHRMGSTRALPGTLJGGGSDSI VPWKRQDEQLSLQVETLYLLDSFFPTLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYQAHAPLLACHSVDGAVPRPSSSXLWGVPTPSPSAVLLST							WDNNNRSTGGSWNFGPQDSNDNKWGEGNK
SSTGSEVEGÖSTGSNIHKAÖSSDSÄINSGARSY RPTHPDCQAVLQTILSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPREGKSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS AATGWKDFKNTGGWNDYKNNNSSNWGGR PDEKTPSSWRENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKKS, PAWNETGRQPNSWNKQHQQQPPQQPPPQP PEASGSWGGPPPPFGNVRFSNSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVILWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTOSQGASSINSASWGQG GKKQMKCSLKGGNDSWMPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MANLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYPEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQUAMLSQLPQDFQFQLACQL LLQQQQQQQLQNQRKISIQAVRQQQEQQLA RMVSALQQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTCKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFRQWTSMME GLFSVATQEAMHRKNGAVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLPASSPPTNK IGSKSSNASWPPPPQFPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTQUDGSTLRTICMQFPLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLAEF ATDDEVSRFLAQAGPPTPAATPSAPAAGWQS LETGQNQSPPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PRIMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSQVETLYLDSFAVILLSPTFLAPVDS PQDCGFKDHQPLTLQALTVELARWTLMLLS TAMYGAHAPLLACHTVDGKVPFPSSAVLLT LEITKLLLCAFSLLVWQGWPQOPPPWRQAA	1						MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN
RPTHPDCQA'U_OTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEVPPEGKSDKGTEGWES AATQTRNSGGWGBASQSNQMKSGWGELS AATQTRNSGGWGBASQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNISSNWGGGR PDEKTTSSWREPPSKDQGWGGGRQPPQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNALSASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQPPQQPPPQP PEASGSWGGPPPPPFGNVRFNSNSWSGGPA TPKDEEPSGWEEPSPGSUSFKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSHNSASWGGG GKKQMKCSLKGGNNDSWMNPLAKÇFSNNG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDPNDIMRKDRSGFRPPNSCLMGTTDS GPYFEKGGSGHG-FGNSTAQSRGLHTPVQPLN SSPSLRAQVPPOFISPQVSASMLKOFFNSGLSF GLFNVGPQLSPQQALDLORKISQAVRQQEQQLA RMVSALQQQQQQLLONQKRISQAVRQQEQQLA RMVSALQQQQQQLLONGKRISQAVRQQEQQLA RMVSALQQQQQQLCPGMKRISGAVRGGSPY NQFDIIPGDTLGGHTQPAGDSWLPAKSPFTVK IGSKSSNASWPPEPGPGVPKKGINDPESDP YVTFGSVLGGTATSPTVDTHOLLINTTGS NSSLNTSLPSPGAWPYSASDNSTTNVHSTSAK PPDVKSTVSPDPJGHNFLSNKMWKNHSIS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRSYWLVLHN LTQLIDGSTLRTICMQHGPLLTFHINLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLAGF ATDDEVSRFLAQAGPTAATSPAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWCVPTVED PHRMGSPAPLPCDLLOGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLLDSFFFFLAPWD PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLACRIVDGKVPPRPSSAVLLT LTPLINGLAGASLLTVELARWTLMLLLS TAMYGAHAPLLACRIVDGKVPPRPSSAVLLT	1				}		SSTGAWDNQKGHPLLENQGNAQAPCWGRSS
QTQIKQDTVWDIEEVPRPEGKSDKGTEG WES AATQTKNSGGWGDAPQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGR PDEKTPSSWNENPSKDQGWGGRQPNQGWS SGKNGWGEEVPOTKNSNWESGASRPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQQPPQQPPPQP PEASGSWGGPPPPGWYRPSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISKMDIDGTSAWG DPNSYNYKNNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSHNSAWGQG GKKQMKCSLKGGNNDSWMPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKRA MILGDPDIMKRDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVQPQLSPQQLALNLQQQQQQLA RMYSALQQQQQQQLQMMSLSQLPQIPGFQLACQL LLQQQQQQLLQNORRISGAVRQQQEQQLA RMYSALQQQQQQQRGMKHSPSIPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMYGMVGGKEAGTESFRKQWTSMME GLFSVATQEANMHINGAIVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPPGPGVPWKGIGMIPDESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTFLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSVNLVLIN LTPQDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLLAEF ATDDEVSRLAQAGGSVRPSVNLVLIN LTPQDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLLAEF ATDDEVSRLAQAGGSVRPSVNLVLIN LTPQDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLLAEF ATDDEVSRLAQAGGSVRPSVNLVLIN LTPQDGSTLRTICMQHGPPLTFPLANTPSAAGWGS LETGQNQSDPVOPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDGI LPSGNQSDPVOPSLALLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDGI LPSGNQSDPVOPSLALLFGGTGLGQWSSSA LPSGNQSGNAGASLWGPPTYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDGI LPSGNQSSACLLDSFFFLAPWDS PQDCGFKDHOPLTLQALTVELLARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFPRSSXACLLDSFFFLAPWDS PQDCGFKDHOPLTLQALTVELLARWTLMLLLS TAMYGAHAPLLALCHVDGRVPPPRSVALLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY
ASTUKNSGGWODAPSOSNOMISGSWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EGGORIGTWGMGASLASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQPPQQPPPQ PEASGSWGGPPPPPGNVRFSNSSWSGPPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTIGSQGSASSHNSASWGQ GKKQMKCSLKGGNNDSWMPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDDDKAT MNLGDPNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQLN SSPSLRAQVPPQFISSQVSASMLKQFPNSCLSP GLFNVGPQLSPQQQQQRQPGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPPGYGSGF SSGGMDYGMVGGKAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPTINK IGSKSSNASWPEPEPQGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTIDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLRPPPGLTNRFSSPWSSTAPRSVRG WGTQDGSRLASASTWSDGGSVRPSVNLVLHN LTPQDGSTLRTICMQHGPLLTHLINLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGGWSSSA GGSSGADLAGASLWGPPTNSSSSLWGVPTVED PHRMGSPAPLLPGDLLOGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTIF LPSSLBPFLQVDSSSSACTLDSFTPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG
ASTEWLOPKNITGGWNDYKNNISSNWGGGR PDEKTPSSWNEIPSDQWGGGRGPNQGWS SGKNGWGEEVDQTKNSNWESSASKPV3GWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQQPPQPPPPQ PEASGSWGGPPPPPGNVEPSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNNIWDENSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSINSASWGQG GKKQMKCSLKGGNDSWMNPLAKOPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDPNDIMKRDSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAGSRGLHTPVQPLN SSPSLRAQVPPPG1SPQVSASMLKQPPNSCLSP GLFNVGPQLSPQQLAMLSQLPQPQFQLACQL LLQQQQQQLLQNQRKJGKAYRQQQQQLA RMVSALQQQQQQQLQNGRGKHSSPIPVGFK PHLDNMVPNALNVGLPDLOTKGPIPGYGSGF SSGGMDYGMVGGKACATESPRKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NOFDIIEGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPPGGVPWKGIGNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK PFDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLTNFKPSSPWSSTAPRSVRG WGTQDSRLASASTWOGGSVRPSVNLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQBAKAQTALHMCVLGNTTILAEF ATDDEVSRLAGAGTALTHCVLGNTTILAEF ATDGEVGRAGAGTALTHCALTTCHTCHTTTATTCHTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTTATTCHTT	1	1					QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES
PDEKTPSSWIENPSKDQGWGGGRQPNQGWG SGKNGWGEVDQTKNSNWESSASKPVSGWG EGQQNEIGTWGNGGNASLASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQQPPQPPPP PEASGSWGGPPPPGNYRPSNSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGAPREPNLPTP MTSKASDSKSMDGWGESDGPVTGARIPS WEEEDGGWNTTTOSQGGASSHNSASWGQG GKKQMKCSLKGGNNDSWMPPLAKQPSNWG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKRA MNLGDFNDIMRKDRSGFPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSCLSP GLINVGPQLSPQLAMLSQLPQIPQFQLACQL LLQQQQQULQNQRISQAVRQQQEQQLA RMVSALQQQQQQQQRPGMKHSPSIPVGFK PHLDMNVPNALNVGLPDLQTKGPPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIPGDTL GGFTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQFQVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTTNVHSTSAK FPDYKSTWSPPPIGHPFTH.SNKMWKNHISS RNTTPLPRPPGLTNPKPSSPWSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAFF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLQGWSSA GGSSGADLAGASLWGPPNYSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSFAVIHLSPIT LPPSSLPFFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWITMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVILLT TAMYGAHAPLLALCHVDGRVPFRPSSAVILLT	1						AATQTKNSGGWGDAPSQSNQMKSGWGELS
SGKNGWGEFVDQTKNSNWESSASKPYSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQPPQQPPPPQ PEASGSWGPPPPPGNVRPNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGAPREPNLPTP MTSKSASDSKSMQDGWGESDQPVTGARHPS WEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKRA MNLGDFNDIMRKDRSGFPPPNSCMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQV3SMLKQFPNSGLSP GLFNVGPQLSFQQLAMLSQLPQIPPQFQLACQL LLQQQQQQQQQQQQQQQQQQMKHSPSPHPVGPK PHLDNMVPNALNVGLPDLQTKGPPDYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVWKGIQNIDPESDP VVTPGSVLGGTATSPIVDTHDQLTRDTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMVKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LLTQLDGSTLRTICMQHGFLLTFHLNLTQGTA LLRYSTKQEAAKAQTALHMCVLGNTTLAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSA GGSSGADLAGASLWGPPNYSSLWGVSTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPFFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT	1						ASTEWKDPKNTGGWNDYKNNNSSNWGGGR
EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQPPQPPPPP PEASGSWGGPPPPPPGNYRFSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGWNTTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKRA MNLGDFNDIMRKDRSGFRPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQLAMLSQLPQIPQFGLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQFGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLOTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESFFKQWTSMME GLFSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPTTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLFSPGAWPYSASDNSTTNVHSTSAK FPDYKSTWSPDPIGHNFTH.SNKMWKNHISS RNTTPLPRPPFGLTNPKPSSPWSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLLAEF ATDDEVSRFLAQAQPPTPATTSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGGWSSAA GGSSGADLAGASLWGFPNYSSSLWGVFTVED PHRMGSPAPLLPGDLLGGGSDS1 1049 2399 A 8748 200 1387 VPWKRQDEGLSLQVETLYLDSPAVIHLLSPTF LPPSSLPFFLQIVDSSSSACTLDSFPFFLAPWDS PQDCGFKDHQPLTLQALTVELARWTIMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT	1			1			PDEKTPSSWNENPSKDQGWGGGRQPNQGWS
PAWNETGROPNSWIKQHQQQOPPQPPPO PEASGSWGGPPPPPPGINVRPSNSWSSGPQPA TPKDEEPSGWEEPSPQSISKKMDIDDGTSAWG DPNSYNYKNVNLWKNSQGGPAPREPNLPTP MTSKSASDKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGRNDSWMNPLAKQFSNMG LLSQTEDNPSKMDLSYGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPGISPQVSASMLKQPPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNGRISQAWLQQEQQLA RMVSALQQQQQQQRQPGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHINGATVAFOKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPGGYPWKGJQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGPTVVED PHRMGSPAPLLFGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILSPTF LPPSSLPPELQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMILLS TAMYGAHAPLLALCHVDGRVPFPRSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1	1					
PEASGSWGGPPPPPGNVRFSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMPPLAKQPSNMG LLSQTEDMPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYYEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQQLQNQRKISPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSSTNVHSTSAK FPDYKSTWSPPFIGHNPTHLSNKMVKNHISS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQDGSTLATICMQHGPLLTHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDYOPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHHLSPTF LPPSSLPPFLQTVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFPPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						EGGQNEIGTWGNGGNASLASKGGWEDCKRS
TPKDEEPSGWEEPSQSISRKMDIDDGTSAWG DPNSYNYKNYNLWDKNSQGGPAPREPNLPTP MTSKSASDKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWQQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQQQLA RMVSALQQQQQQQQQQRGMKHSPSHPVGPK PHLDNMVYNALNVGLPDLQTTGCPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTIDQLLRDNTTIGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTTPLPRPPGGLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTTQALTYCLARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	ļ						PAWNETGRQPNSWNKQHQQQQPPQQPPPPQ
DPNSYNYKNYNI, WDKNSQGGPAPREPNI, PTP MTSKSASDSKSMQDGWGESDGVTGARHPS WEEEEDGGVWNTTGSQGSASSINSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKNDLSVGSLSDKKFDVDKRA MMI, GDFPDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPOFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQPPOFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRGMKHSPSHPVGFK PHLDNMVPALNVGLPDLQTKGPIFGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQFGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDVKSTWSPDPIGHNFTHLSNKMWKNHISS RNTTPLPRPPGLTNFKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LIPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGGNQSPYGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDGDSLQVETLYLDSPAVHILSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTTQALTYLDSPAVHILLSTF LPPSSLPFFLQIVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTYLDSPAVHILLSTF LPPSSLPFFLQIVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTYLDSPAVHILLSTT LPPSSLPFFLQIVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTYLDSPAVHILLSTT LPPSSLPFFLQIVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTYLDSPAVHILLST TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							PEASGSWGGPPPPPPGNVRPSNSSWSSGPQPA
MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQQLQNQRKISQAVRQQQEQQLA RMYSALQQQQQQQQQQQRGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAKAQTALHMCVLGNTTILAEF ATDDEVSFLAQAQPPTPAATTSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGYPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPPFRSSAVLLT LTAMYGAHAPLLALCHVDGRVPPFRSSAVLLT ELITKLLLCAFSLLVGWQAWPQGPPPWRQAA	ļ						
WEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQFGMKHSPSHPVGFK PHLDNNVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPPDPIGHNPTHLSNKMWKNHISS RNTTPLPPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATTSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSFAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPITLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSYGSLSDKKFDVDKKA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFFNSGLSP GLFNVGPQLSPQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMOHGPLLTFHLNLTQGTA LIRYSTKQEAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTTQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1 :						
LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQVSASMLKQFPNSGLSP GLFNVGPQLSPQVSASMLKQFPNSGLSP GLFNVGPQLSPQVSASMLKQFPNSGLSP GLFNVGPQLSPQVSASMLKQFPNSGLSP GLFNVGPQLSPQVSASMLKQFPNSGLSP GLFNVGPQLSPQVSASMLKQFPNSGLSP GLFNVGPQLSPQVATGENDEY PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTIQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LIPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPPRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA						,	GKKQMKCSLKGGNNDSWMNPLAKQFSNMG
GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLQNQRISQAVRQQQEQQLA RMVSALQQQQQQQQRPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESFFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFTPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPPRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						[*]
SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPOPQLACQL LLQQQQQQLLQNQRKISQAVRQQQQQQLAA RMVSALQQQQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQVVDSSSSACTLDSFPFPLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQFQVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	i				1		
LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPTF LPPSSLPPFLQVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVILLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
RMVSALQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLCAFSLLVGWQAWPQGPPPWRQAA							
PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHILSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHIN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LIPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1			i -	•		
YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					İ	l	
NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LIPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					•]	l	
FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA]				ļ	į	
RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					{	1	
WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA			i				
LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA]					ĺ	
LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA			ļ		ļ		
ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA]						
LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					ļ		
GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA			İ				
PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA			ĺ		1	. 1	
1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1040	2200	<u> </u>	0740		1202	
PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1049	2399	A	8748	200	1387	
TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
L	.			1			
				1			PFALSALLYGANNNLVIYLQRYMDPSTYQVL

uence unce unce unce unce unce unce unce	SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
uence 99496 gortespondi go to fist a mino - acid residue of peptide residue of peptide sequence pep							
amīno acid residue of sequence peptide s	seq-		i	09/496			
Peptide of peptide sequence	uence			914			
Peptide	1		ĺ	[
			<u> </u>			sequence	
SNIKIGSTAVLYCLCERRISYROGGALILL MAGGCYAGGGLOYOFOTIPSPPYAAASS MPLHTPIGILLILILYCLISGISSVTEILMKK QRLPIALQNIELYTCYVOLING, GHAGGGOP GILEGFSGWAALVVLSQALNGLLMSAVMKH GSFTRIATILLIGIAMRLYYGGR GILEGFSGWAALVVLSQALNGLLMSAVMKH GSFTRIATILLIGIAMRLYYGGR FRIATILLIGIAMRLYYGGR FRIATILLIGIAMRLYYGGR FRIATILLIGIAMRLYYGGR WVSSMGFEILEGVGGGPPOLRNYALLALP RTULPLHFILPIFLAVPAHRCALPGAPANFS RTULPLHFILPIFLAVPAHRCALPGAPANFS RTULPLHFILPIFLAVVAGALVATORYST A FISS SISSINGT GOV AVAGYLSGROFFOR A FISS SISSINGT GOV A	i		ł				
MAGGACYAAGGLQYPONILESPPTAAAASP MPHHPIGLILLII, YCLISGISSVYELLIMKR QRLPLALQNLFL YFFGVLINL, GHAGGGSQ GLEGFSQWALVV.SQAINGLIASAMKH GSSTRLFVVSCSLVVNAVLSAMLRLQLTAA FFLATLIGLAMBLLYGGSR		-	 		Sequence		
MPI-HITPLGILLILLYCLISGLSSYTTELLMRG QR.P.LA.QN.F.P.YTGVLINLGLHAGGGSGP GLEGFSGWAALVVLSQALNGLMSAYWKH GSTRIF-VYCSCSLVVNAVISAVLHRQLTAA GSTRIF-VYCSCSLVVNAVISAVLHRQLTAA FFLATLLIGLAMRLYYGSR	1		}		Ť		MAAGACYAAGGLOVPGNTLPSPPPAAAASP
1050 2400 A 8758 3 1660 WYSSMGFEELLEGY (GSTRIPY OSCI, VNANJA, SAVLIRI QLTAA FFLATILIGI LAKRIL YYGSR FFLATILIGI LAKRIL YYGSR WYSSMGFEELLEGY (GSTGFEP) CRIVALILAD RVILPIHFILPIFLAAVPAHRCALPGAPANFS HQDVWILAHLPREPOGTISSCLRFAYPQALI NTILGERGS (GELEDEPATVPCS GGWEY) DH SEFSSTIATES (WYDLVCE) (KGLNRAASTFFFA GVLVGAVAFGYLSDERFGKREILLVAYVSTILV LGLASAASVSYWFAITRILTGSALAGFTIU MPLELEWLDWEHTVAGVLSTFWTGGVML LALVGYLRBWRWILLIAVTLFCAPGISS WW VPESARWILTGGVIKEAHRYLHGARINGR PVCEDSFS (GARWILTGGVIKEAHRYLHGARINGR PVCEDSFS (GARWILTGGVIKEAHRYLHGARINGR PVCEDSFS (GARWILTGGVIKEAHRYLHGARINGR PVCEDSFS (GARWILTGGVIKEAHRYLHGARINGR PVCEDSFS (GARWILTGGVIKEAHRYLHGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR PVCEDSFS (GARWILTGGARINGR) PVCEDSFS (GARWINGS) PVCEDSFS (GARWILTGARINGR) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSFS (GARWINGS) PVCEDSF (GARWI			İ				MPLHITPLGLLLLILYCLISGLSSVYTELLMKR
1050 2400 A 8758 3 1660 WVSSMGFEELLEQVGGGFGPQLRIVALLAL							
PFLATILIGI.AMRLYYGSR							•
1050	1		1				\
RVILPHTLIPIFIAAVPAIRCALPGAPANES HQDVWLEAHLPREPDGTISSCLRFAYPQALP NTTLGERQSRGELEDEATYPCSQGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFA GVLVGAVAFGYLSDFGFGRETUAYYSTLUAYYSTLU LGLASAASVSYVMFAITRTLTGSALAGFTIIV MPLELEWLDVPHRIVAGVLAGALAGFTIIV MPLELEWLDVPHRIVAGVLAGALAGFTIIV ALLAUGVLIRDWRWLLLAUTLPCAPGILSILW VPESARWLITQGHYKEAHRYTIGSALAGFTIIV RTPRILRHISLCCVVVWFGVNFYSCLSLDVS GLGLNYQYQTQLLFGAVELPSKLLVYLSVRYA GRRITQAGTILGTALAAGTALLPGSVWLS TVLAVMGKAFSAAFTITAYTYSLLSTPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAAGTALLPGTRAQLPETT QDVERKSAPTSLQEEEMPMKQVQN TCHLAMGSLAFTSLAFTSCLYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LFKLTYGGIALLAAGTALLPGTRQALPETT QDVERKSAPTSLQEEEMPMKQVQN TCHLAMGSLGFVHYSGASPGSEGDEETQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TCHLNSDPLIELCQCPLCJLDCSREQLIAHV YQHTAAVYSAKSYMCPVCGRALSSPGSLGR HLHHSEDQRSNCAVCGARTTSHATTNSEKLP EVLNMESLPTVHNEGPSSAGKOLAFSSPVYT AGGLLVCNNCAAYFKLLEAQTFSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERR QNEPLEVRLQRLERERTAKKSRRDNETPEERR VRRMRDREAKRLQRNOFTDEGRABRALKAR LEKMDMMLRAQFGQDPSAMAALAAEMNFP QLPYSGVELDSQLLGKMAFERRAKLQRND EAMRLKRANETPEERQALIDSGNSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYFGGGA ASLVFVFTRRSGPSGTASVAAMAYHSCYGA HGSKHRARAAPDPPLFDTSGYSSQFGGY PATGADVAFSVNELLGUNGSVSLH ASHGKDMYHKERTYSVSLKLYVFRAVDTA YVAKKLGLLVFPYTHONWEVQYSRDAPLPP RQDLNAPDLYPTMATITYVLLAGMALGIQK RFSPEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHALAYSGYKYVGMULSVLTIGL LPGSDGGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQELITY WITTHLVR 1053 2403 A 8768 2 712 RPFRVWYPELRELSAAAPRWSHRTAPGIMVF YFTISSSNYSSAYTIYMGGRVYFNEDLIRHGW PEDIWFFHVDKLSSAHVYLRLHKGENEDPKE VLMDACHLVKANSIQGCKMVNIVVYTFW SALKKLADMDVQQJGHRQKCRVIVTVEKK VNEILNRLEKTKVERFPDLAAEKBERDREER NEKKAQQQEMKKREKEMKKKREMDELRSY SSLIMKVERMSSNQDGNSDEFM RKKAQQQEMKKREKEMKKKREMDELRSY SSLIMKVERMSSNQDGNSDEFM	1050	2400	Δ	8758	3	1660	
HQDVWLEAHLPREPDGTLSSCLRFAYPQALP NTIGEREQSRGLEDERATYCOGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFA GYLVGAVAFGYLSDRFGRRELLLVAYVSTLV LGLASAASVSYVMFAITRILTGSALAGFTIIV MPHELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLCAPGILSLWW VPESARWLITQGHYKEAHRYLLHCARLINGR PVCEDSFSQEAVSKVAAGERVYRRSYLDLF RTPRLRHISLCCVVWFGVNFSYYGLSLDVS GLGLNYYGTQLHGAVELPSKYVRRSYLDLF RTPRLRHISLCCVVWFGVNFSYYGLSLDVS GLGLNYYGTQLHGAVELPSKYVRRSYLDLF RTPRLRHISLCCVVWFGVNFSYYGLSLDVS GLGLNYYGTQLHGAVELPSKYVRRSYLDLF RTPRLRHISLCCVVWFGVNFSYYGLSLDVS GLGLNYYGTQLHGAVELPSTYLVLR GRRLITQAGTLLGTALAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEBMPMKGVQN TULAVMGKAFSAAPTTAVLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN ABTOLICATION GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPMKGVQN GREEN TOMANSASTSLQEEBMPM GREEN TOMANSASTSLQEEBMPM GREEN TOMANSASTSLQEEBMPM GREEN TOMANSASTSLQEEBMPM GREEN TOMANSASTSLQEBMP GREEN TOMANSA	1000	2100	**	0,50		1000	
SEFSSTIATESQWDLVCEQKGLNRAASTFFRA GVLVGAVAFGVLSDRFQLLVAYVSTLV LGLASAASVSYVMFATTETT.TGSALAGFTIIV MPLELEWLDVEHRIVAGVLSSTWTTGGVML LALVGYLIRDWRWLLLAVTLPCAPGILSI.W VPESARWLLTQGHVKEAHRYLLHCARINGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISL.CCVVVWFGVNFSYYGSLDVS GLGLNVYOTQLFGAVERSVLAVISVRYA GRRI.TQAGTILGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETT QDVERKSAFTSLQEELBWFMKQVQN 1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSGGDEETTJOEVSSHTS CEDGGVVKVKEKLENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSRSQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLHISEDQRSNCAVCGARTSHATFNSEKLP EVLNMESLPTVINGFSSABGKDLAFSPVVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDAFSPVVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDETPEERE VRRNDREARKICRNGPTDEQRARRLQRDR EAMMLRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDFSAMAALAAEMNFF VRRNDREARKICRNGPTDEQRARRLQRDR EAMMLRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDFSAMAALAAEMNFF VRRNDREARKICRNGPTDEQRARRLQRDR EAMMLRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDFSAMAALAAEMNFF VRRNDREARSVINELLGGMAAFGALGRYBGGGG PATGADVAFSVINELLGGMANAVMAYGSSI ASIGKDMVHKELHFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRAPLPLP RQDLNAPDLYPTTMAFTIYVLLAGMALGIQK RSSPEVLGLCASTALVWVVMEVALLLIGLYL ATVRSDLSTFFILLAYSGVYTYGMLSVLTIGL LFGSDGYYVALAWTSSALMYFFVSSLRTAAL GPDSMGGPVPRQRLQLYLTLIGAAAFQPLITY WLTFHLVR VARKLGLLVFPYTHQNWEVQYSRAPLING PDSMGGPVPRQRLQLYLTLIGAAAFQPLITY WLTFHLVR VARKLGLLVFPYTHQNWEVQYSRAPLPLP RQDNAMPTVSLLLIGGVTATAAL GPDSMGGPVPRQRLQLYLTLIGATAAAFQPLITY VAKKLGLLVFPYTHQNWEVQYSRAPLPLP RQDNAMPTVSLLGLAGAAFQRLIRER VRRNDRAGGROGGGFRQKGCGGARNGNVVVVTVEVL VNECALLLIGLYL ATVRSDLSTFFILLAYSGVYTYGMCKYTYGKK VNECALLAGRACGCROREE NEKKAQIQEMKREKEEMKKRRMDELRSY SSILKKVENNSNQDGNDSDEFM NEKKAQIQEMKREKEEMKKRRMDELRSY SSLMKVENNSNQDGNDSDEFM NEKKAQIQEMKREKEEMKKRRMDELRSY SSLMKVENNSNQDGNDSDEFM							· .
GVL/GAVAFÖYLSDRFGRRILLIVAYVSTLU LGLASAASVSYVMFATITGSALAGFTIIV MPLELEWLDVEHRIVAGVLSSTFWTGGVML LALVGYLRDWRWLLLAVTLPCARGILSLWW VPESARWILTQGHVKEAHRYLLHCARLNGR PVCEDSFSQEAVSKVVAGGRVVRRSYYGLSLDVS GLGLNVYOTQLLFGAVELPSKILVYLSVRYA GRRLTQAGTLLGTALAFGTRILLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLTSSLLYFTVLR QTGMGLTALVGRILLAGTALLVSSDMKSWS TVLAVMGKAFSEAAFTTATLLTSSLLYFTVLR QTGMGLTALVGRILLAGTATLVSSDLYFTUL QDVERKSAFTSLQEEMPMKQVQN TGRLTNSSPESLYFTUL QDVERKSAFTSLQEEMPMKQVQN TGRLTNSSPESLYFTUL QDVERKSAFTSLQEEMPMKQVQN TGRLTNSSPESLAFT TGRLTNSSPESLAFT TGRLTNSSPLIFT TGRLT	1						NTTLGEERQSRGELEDEPATVPCSQGWEYDH
LGLASAASVSYMMFAITRITLTGSALAGFTIIV MPILELEWLDVEHRTVACSISTEWTGGVM. LALVGYLIRDWRWLLLAVTLPCAPGILSLWW VPESARWLLTQGHVKEAHRYLLHCARLNGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRIRHISLCCVVVWFGVNFSYYGSLDVS GLGLNYVTGULFGAVELPSKILVYLSVRYA GRRITQAGTILIGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTLALTFSELYPTUR QTGMGLTALVGRLOGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQALPETI QTGMGLTALVGRLOGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQALPETI QTGMGLTALVGRLOGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQALPETI QTGMGLTALVGRLOGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQALPETI QTGMGLTALVGRLOGSLAPLAGLEDGVWLS LPKLTYGGIALLAAGTALLLPETRQALPETI LPKLTYGGIALLAAGTALLLPETRQALPETI TGNLNSDPLLELCQCPLCQLDCGSRSQLIAHV YQHTAAVVSKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSRSQLIAHV YQHTAAVVSAKSYMCPUGPVGGNEVSHEHV TGNLNSDPLLELCQCPLCQLDCGSRSQLIAHV YQHTAAVVSAKSYMCPUGPVGGNEVSHEHV TGNLNSDPLLELCQCPLCQLDCGSRSQLIAHV YQHTAAVVSAKSYMCPUGPVGGNEVSHEHV TGNLNSDPLLELCQCPLCQLDCGSRSQLIAHV YQHTAAVVSAKSYMCPUGPVGGNEVSHEHV TGNLNSDPLLELCQCPLCQLDCGSRSQLIAHV YQHTAAVVSAKSYMCPUGPVSGNALSPSGLGR HLIHSEDQRSNCAVCGARTISHATTNSEKLP EVLNMBSLETTVIRNEGFSSAEGKDLAFSPVSLAF AGILVCNNCAAYRKLLEAGTDFSVSLAF VRNRDREARKICRMGDTBCRARRLCRDD EAMBLKRANSTPEKRQARLIRERFAKRLKRR LEKMDMMLRAOFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASSLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSOFGGY PATGADVARSSVIBLLGDPMANVAMAYGSSI ASSIGKDMYHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHONIEVGVSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RSPEVLGLCASTALVWVVMEVLALLIGLYL ATVRSDLSTFFILLAYSGYKYYGMLSVLTGL LGCSDGYYYALAWTSSALMYFIVSLATALL GPDSMGGPPYRQRLQLYTTLAGMALGIQK RSPEVLGLCASTALVWVVMEVLALLIGLYL ATVRSDLSTFFILLAYSGYKYYGMLSVLTGL LGCSDGYYYALAWTSSALMYFIVSLKTAAL GPDSMGGPPYRQRLQLYTTLAGAAFQFLIDVF YFTSSSVNSSAYTTYMGKDKYSTYGGMTVF YFTSSSVNSSAYTTYMGKDKYSTYGGMTVF YFTSSSVNSSAYTTYMGKDKYSTGEDDFKE VMDCAHLVKANSIQQCGHRACKAEGERBER NEKKAQIQEMKREKEEMKKRRMDELRSY SSLMKYBNSSNQDGNDSDEFM NEKKAQIQEMKREKEEMKKRRMDELRSY SSLMKYBNSSNQDGNDSDEFM							
MPLELEWI,DVEHRITVAGVLSSTFWTGGVML LALVGYLRDWRWILLAVTLPCAPGILSLWW VPESARWILTQGHYKEAHRYLHCARINGR PVCEDSFSQEAVSKVAAGERVYLHCARINGR PVCEDSFSQEAVSKVAAGERVYLHCARINGR PVCEDSFSQEAVSKVAAGERVYLHCARINGR PVCEDSFSQEAVSKVAAGERVYLHCARINGR PVCEDSFSQEAVSKVAAGERVYLHCARINGR PVCEDSFSQEAVSKVAAGERVYLHCARINGRYPYCILDFS GIGINVYQTQLIFGAVELPSKILLVYLSVRVA GRRITQAGTILIGTALLVSSNKSWS GRITQAGTILLGUNGTILLVSSNKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGITALVGRIGGSIAPLAALLDGVNL; LPKLTYGGIALLAGTAFLLVSSDKKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGITALVGRIGGSIAPLAALLDGVNL; LPKLTYGGIALLAGTAFLLPFTRQAQLPETI QDVERKSAPTSI.QEEEMPMKQVQN ERIPTVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVEHEV TGNLNSDPILELCQCPLCQLDCGSREQLIAHV YQHTAAVVSASSYMCPVCGRALSSPGSLGR HILHSEDQRSNCAVCGARITSHATFNSEKLP EVINMESIPTVHNEGPSSAEGKDIAFSPGVTP AGILLVCNNCAAYRKILEAQTPSVREWALRR QNEPLEVRLQRIERERKAKRRANETPERS VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARILREREAKRLKRR QNEPLEVRLQRIERERKAKSRRDNSTPERER VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARILREREAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLKRR QNEPLEVRLQRIERERAKRLRR QNEPLEVRLQRIERERAKRLRR QNEPLEVRLQRIERERAKRLR QNEPLEVRLQRIERERAKRLR QNEPLEVRLQRIERERAKRLR QNEMCH QNEPLEVRLQRIERERAKRLR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEPLEVRLQRIERERAKR QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QNEMCH QN						1	
LALVGYLIRDWRWLLLAVTLPCAPGILSLWW VPESARWLLTQGHVKEAHGYLIHCARLNGR PVCEDSPSQEAVSKVAAGERVVRRPSYLDLF RTPRLRIISLCCVVWPGVNFSYGLSLDVS GLGINVYQTQLLFGAVELPSKLLVYLSVRYA GRRITQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAVFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEBMFMQVQN 1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGEEFTTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSPPLLELQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSFGSLGG HLLHSEDGRSNCAVCARTHATHNSEKLP EVLNMESLPTVHNEGPSSAEGKDLAFSPPVYP AGILLVCNNCAAYKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMEDREAKRLQRINGFTDEGRSRLQRDR EAMFILKARNETPERQARLIREREAKRLKRR LEKMDMALRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEQNSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRYARPGGIGRYPGRGAA ASLVFVPTRRSGPSGTASVAMAYHSGYGA HGSKHRARAAPDPPFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTPMAFITYVLLAGMLGJGK RFSFEVLGLCASTALVWVMEVLALLIGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LPGSDGYYVALAWTSSALMYFIVVSLTAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR PRPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSSAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYYJRGKDKYENEDLIKHGW PEDIWFHYDKLSKAHYKREMBELRSY SILKKYENINGKYENEDLIKHGENEDLIKHGW PEDIWFHYDKLSKANGWYFSRCSLGACKPTVCSMP SILKKYANGWYFSRCSLGACKPTV					•		,
Pyesarwiltoghykeahrylihcarinog Pycedspsqayskyagervyrpsyldle RTPRIRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNYYGTQLIFGAVELPSKILVYISSWRSWS GRRITQAGTILGTALVERISSDMKSWS GRRITQAGTILGTALVERISSDMKSWS TVIAVMGKAFSEAAFTTAYLFTSELYPTVIR GTGMGLTALVGRIGAGPLAALIDGVWLS LPKLTYGGIALLAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEMPMKQVQN EIRTPVAVSSAPSGDSEGDEEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNINSPPLIELCQCPLOLGCSREQLIAHV YQHTAAVVSARSYMCPVCGRALSSPGSLGR HILHISEQRSNCAVCGARFTSHATTNSEKLI EVINMESLFTVHNEGFSSAEGKDIAFSPPVTP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERKAKSRRQNETPERE VRRMEDREAKRLQRMGETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKNDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH ASLVFVFTRRSGPSGTASVAAMAYHSGYGA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA AGSKHRARAAPDPPPI-FDDTSGGYSSQPGGY PATGADV AFSVNELLDEPMANVAMAYGSSI ASHGKDMYHKELHRFVSVSKLKKFFAVDTA YVAKKLGLVPPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTJMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWMEVLALLIGJVL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYJTRSLRTAAL GPDSMGGPVPRQRLQLCASTALVWMEVLALLIGJVL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYJTRSLRTAAL GPDSMGGPVPRQRLQLJCLASTALVYMEVLALLIGJVL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYJTRSLRTAAL GPDSMGGPVPRQRLQLJCLASTALVYMEVLALLIGJVL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYJTRSLRTAAL GPDSMGGPVPRQRLQLJCASTALVYMEVLALLIGJVL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYJTRSLRTAAL GPDSMGGPVPRQRLQLJCASTALVYMEVLALLIGJVL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYJTRSLRTAAL GPDSMGGPVPRQRLQLJCASTALVYMEVLALLIGJVL ATVRSDLSTFHLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYJTRSLRTAAL GPDSMGGGVPRQRQLJCAAACHONNONVYTYTW SNLKKTADMDVQQGFHRQNDVKTVTVEKK VLMDCAHLVKANSIQGCKMNNNVVYTTPW SNLKKTADMDVQQGFHRQNDVKTVTVEKK VNELNRLEKTKVERFDLAAFKEGUEDFME NEKKAQQQEMKKREKEMKKREMDELRSY SSLMKVENNSNQDGNDSDEFM							_
PVCEDSPSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGISLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTATAYLFTSELYPTVLR QTGMGLTALVGRLQGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLDFTRQAQLPETI QDVERKSAPTSLQEEMPMKQVQN 1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEETITQDEVSSHTS EEDGGVVKVEKLEURGPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLLAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHISEDQRSNCAVCGARPTSHATFNSEKLP EVINMESLPTVINNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYFILLEQTTSYKRWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEGQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARFGGLGRYPGRGAA ASLVFVFTRRSGPSGTASVAMAYHASGYGA HGSKHRARAAPPPFLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTPTMAFITYVLAGMALGIGK RFSEPVLGLASTALVWVVMEVLALLLGLTL ATVRSDLSTFHILAYSGYKVYGMI.SVLTGL LFGSBGYVALA WTSSALMYFIVRSLTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAFGIMVF YFTSSSVNSSAYTTYMGKDKVENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTTPW SNLKKTADMDVGQIGFHRGROVKVIVTVEKK VNELNIBLEKTKVERFFDLAAEKSCRDEER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKFTVCSMP							
GIGLNYVOTOLLEGAVELPSKILLVYLSVRYA GRRLTQAGTILGTALAFGTILVSSDMKSWS TVLAVMGKAFSEAPTTAYLFTSELTYPTVLR OTGMGLTALVGRILGSLAPLAALLDGVWLS LPKLTYGGIALLAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEMPMKQVQN 1051 2401 A 8759 515 1625 EERTPVAVSSAPSGDSEGDEEPTTQDEVSSHTS EEDGGVVKVEKLEDTGQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLHISEDQRSRCAAVCGARPTSHATFNSEKLP EVLNMESLPTVINEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETTEREE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKKR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH LEKMDMMLRAQFGDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLQDPMANVAMYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYPTTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVWEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMLSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYSNEDLIKHGW PEDIWFHVDKLSSALMYFIVRLIKINGENEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRIEDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKRVTRICDIPKE VLMDCAHLVKANSIQGCKMNNNVVYTYPW SNLKKTADMVQQIGFHRKREKEEMBELRSY SSLMKVENMSSNQDGNDSDEFM							
GRRLTQAGTILGTALAFGTRILVSSDMKSWS TVLAVMGKAFSEAAFTTAYIFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGWWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLEEMPMKQVQN 1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEEFTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TONLNSDPLIELCQCPLQQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLIHSEDQRSNCAVCGARPTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILEAQTFSVRKWALRR QNEPILEVRLQNERERTAKKSRRONETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRAGGRVARFGGLGRYPGRGAA ASLVFVPTRRSSPSGTASVAAMAYHSGYGA HGSKHRARAAFDPPLFDTGGGYSSOFGGY PATGADVAFSVHILLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNEVQYSRDALPP RQDLNAPDLVTPTMFTYVLLAGMALGIQK RSFEVLGLCASTALVWVVMEVLALLIGLYL ATVRSDLSTFHLLAYSGYKYVOMILSVLTGL LPGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKKYENEDLIKHGW PEDLWFHVDKLSSAHVYLRLHKGENEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVQQIGFHRQDVKIVTVEKK VNEILAMDDVQQIGFHRQDVKIVTVEKK VNEILAMDDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDDVCVTVEKK VNEILANDMDVQQIGFHRQDVCVTVEKK VNEILANDMDVQQIGFHRQDDVCVTVEKK VNEILANDMDVQQIGFHRQDDVCTVEKK VNEILANDMDVQQIGFHRQDDVCDSDEFM NEKKAQIQEMKKREEBMKKKREDELRSY SSLMKVENMSSNQDGNDSDEFM							RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS
TVLAVMŒKAFSEAAFITAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEMPMKQVQN 1051 2401 A 8759 515 1625 EIRTPVAVSSAPSGDSEGDEEFTTQDEVSSHTS EEDGGVKVEKELENTEQPVGGNEVVEHEV TONLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSASVMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARPTSHATTNSEKLP EVLNMESLPTVHNEGPSSAGKGLIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRIMGETDEQRARRLQRDR EAMRIKRANETPERRQARLIREREAKRIKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLIGKMAFEEQNSSSLH RIGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHISGVGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNFLLLGDPMANVAMAYGSSI ASHGKDMYHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVMEVLALLIGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLITY WLTFHLVR 1053 2403 A 8768 2 712 RPPRWYPELRELSAAAPRWSHRTAPGIMVF YFTSSVMSAYTIYMGKKYEMEDLIKHGW PEDIWFFHDKLSSAHYYLRLHKGENEDIPKE VLMDCABLLVKANSIQGCKMINVNVVYTPW SNLKKTADMDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAKEGCRDREER NEKKAQIQEMKKRRKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACKNSCVYFSRCSLGACKPTVCSMP							
CTGMGLTALVQRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTIS.QEEBMPMKQVQN							
LPKLTYGGIALLAAGTALLLPETRQAQLPETI							
QDVERKSAPTSLQEEEMPMKQVQN							
1051 2401 A 8759 515 1625 EIRTPWAVSSAPSGDSEGDEETTIODEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGRNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFISHATTNISEKLP EVLINNESLPTVHNEGPSSAGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARTLORDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDFSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH RIGHGGRDRRGGGRVARPGGLGRYPCRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYPALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYTLLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFFIVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFFDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SLMKVETAMDVOGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFFDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SLMKVETAMDVSQOGDSDEFM							ODVERKSAPTSI.OFEEMPMKOVON
EEDGGVVK VEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPSGLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLFTVHHEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRNMEDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRORGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASSHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVWEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLKTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADDVQQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFFDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1051	2401	A	8759	515	1625	
VQHTAAVVSAKSYMCPVCGRALSSPGSLGR HILLHISEDQRSNCAVCGARTTSHATIFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTTSVRKWALRR QNEPLEVRLORLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMPLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGRGDRAGGRVARPGGLGRYPGRGAA ASLVFVPTRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYYGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNIVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	-						•
HILIHSEDQRSNCAVCGARFTSHATFNSEKLP EVINMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMILRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVFTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSFEVLGLCASTALVWVVMEVIALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKRREKEEMMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1						
EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKILEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH QLPVSGVELDSQLLGKMAFEEQNSSSLH RIGSKHRARAAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSGI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYPIVRSLRTAAL GPDSMGGPVPPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP							1 7
AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRGGGRVARPGGGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGTYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGGW PEDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP							
QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYTPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGGW PEDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNYNVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1						
VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMYHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLASTALLWVVVMEVLALLLIGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFFHVDKLSSAHVYLKILHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNIEKTKVERFPDLAAEKBCRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP							
EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDDIWFHVDKLSSAHYYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	j					·	
QLPVSGVELDSQLLGKMAFEEQNSSSLH 1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKVFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHILAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYPIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGGCKMNVNVVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	}						EAMRLKRANETPEKRQARLIREREAKRLKRR
1052 2402 A 8763 1106 70 RHGHGGRDRRGGGRVARPGGIGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNYNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP							
ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFFYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1052	2402		9762	1100	70	
HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFFYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1032	2402	A	8/03	1100	70	•
PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			- 1	İ			
ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP							
RQDLNAPDLYIPTMAFİTYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1	}	1	1	}	ł	
RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP				Ì	}	1	
ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		ļ	ļ	1	ļ		• • • • • • • • • • • • • • • • • • • •
LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP]	l	ĺ			
GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR 1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTTYMGKDK YENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKBCRDREER NEKKAQIQEMKKREKEEMKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		1	ļ			1	
WLTFHLVR		ĺ	- 1	İ	1	ĺ	
1053 2403 A 8768 2 712 RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKBCRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		}	1				
YFTSSVNSSAYTTYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP	1053	2403	A	8768	2	712	
VLMDCAHLVKANSIQGCKMNNVNVYYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKBCRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		}	ļ		ļ	ļ	
SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP					İ		
VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		ł					
NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM 1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP		ľ	ĺ		- 1	ľ	
					}		
1054 2404 A 8769 344 527 REATTLACRNSCWVFSRCSLGACKPTVCSMP			ļ		}	. }	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1054	2404	A	8769	344	527	
						**	•

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	f	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
	lana		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	Ì		914			
[ĺ	ĺ	[amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	1		peptide		/=possible nucleotide deletion, \=possible
	1	ł		sequence		nucleotide insertion
1055	2405	Α	8770	430	1104	QQESPAAGAARMNCKEGTDSSCGCRGNDEK
				1		KMLKCVVVGDGAVGKTCLLMSYANDAFPEE
1	i			1		YVPTVFDHYAVTVTVGGKQHLLGLYDTAGO
1			Ì			
j		ŀ	ŀ	l .		EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV
1		1	1	l .	j	QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK
			1	f		TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL
	j l		l	Ì		ECSALTQKGLKAVFDEAILTIFHPKKKKKRCS
			ì			EGHSCCSII
1056	2406	A	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	A	8778	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH
1057	2407	Α	0//0	3	[* ′ ′	
				•		RRDQKWHDKQYKKAHLGTALKANPFGGAS
1	ĺ		j			HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK
]						NGKKITAFVPNDGCLNFIEENDEVLVAGFGR
j l		ŀ	Ì	ļ		KGHAVGDIPGVRFKVVKVANVSLLALYKGK
			l			KERPRS
1058	2408	Ā	8808	171	881	PGLSOEPSGSMETVVIVAIGVLATIFLASFAAL
		**	55.55			VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME
]			ļ	ļ		TQSEPSELELDDVVITNPHIEAILENEDWIEDA
			•			SGLMSHCIAILKICHTLTEKLVAMTMGSGAK
						MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL
1						DPKLLDARTTALLLSVSHLVLVTRNACHLTG
						GLDWIDOSLSAAEEHLEVLREAALASEPDKG
]						LPGPEGFLQEQSAI
1059	2409	A	8809	246	757	MRLOGAIFVLLPHLGPILVWLFTRDHMSGWC
.005	2.05	••	0005	2.0	, , , ,	EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY
<u> </u>						·
]					_	LVWKDLGGGLGWPLALPLGLYAVQLTISWT
i .						VLVLFFTVHNPGLALLHLLLLYGLVVSTALI
						WHPINKLAALLLLPYLAWLTVTSALTYHLWR
						DSLCPVHQPQPTEKSD
1060	2410	Α	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	A	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA
				,		FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF
i i						GAMLNIAAVLCIATIYVRYKQVHALSPEENVI
						IKLNKAGLVLGILSCLGLSIVANFOKTTLFAA
	1					
	İ					HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH
]						GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS
						GNFGTDLEQKLHWNPEDKGYVLHMITTAAE
						WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL
					İ	TLYDTAPCPINNERTRLLSRDI
1062	2412	Ā	8824	1	763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA
				-		GARAPQLCSSWQRRSAPAMSRGLOLLLLSCA
						YSLAPATPEVKVACSEDVDLPCTAPWDPOVP
					'	
				'		YTVSWVKLLEGGEERMETPQEDHLRGQHYH
						QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR
		l			•	CTLQDPDGQRNLSGKVILRVTGCPAQRKEET
						FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI
						FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT
		ļ				ELV
1063	2413	A	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE
		- 1				HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT
		l		ļ		AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ
		l	i			
				į		KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH
	ļ			ļ		CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL
						TR
1064	2414	A	8835	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE
	[į		-	İ	LNKQVSELSQLYKEAQAELEDYRKRKSLEDV
1					j	TAEYIHKAEHEKLMQLTNVSRAKAEDALSE
	.]	ļ	ļ	. }	J	MKSQYSKVLNELTQLKQLVDAQKENSVSITE
		l				HLOVITTLRTAAKEMEEKISNLKEHLASKEVE
				لينــــــــــــــــــــــــــــــــــــ		

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	Į.	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
dence			914	ng to first amino acid		Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	
				sequence		nucleotide insertion
}		 		Soquerico		VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS
i		i				SLESEVSVLASKLKESVKEKEKVHSEVVQIRS
						EVSQVKREKENIQTLLKSKEQEVNELLQKFQ
]			QAQEELAEMKRYSESSSKLEEDKDKKINEMS
}	•					KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
						LQQQVKQLQNQLAECKKQHQEVISVYRMHL
						LYAVQGQMDEDVQKVLKQILTMCKNQSQK
						K
1065	2415	Α	8841	3 .	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
1 .						APLPTGRAQMSPSGRLCLLTIVGLILPTRGQTL
1 :						KDTTSSSSADATIMDIQVPTRAPDAVYTELQP
						TSPTPTWPADETPQPQTQTQQLEGTDGPLVT
		l				DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT
1						DPQTLKPSGFHEDDPFFYDEHTLRKRGLLVA
1066	2416	A	8853	3806	2204	AVLFITGIILTSGKCRQLSRLCRNHCR
1000	2410	Α.	6623	2000	4404	FVGEQEGGCEAGAGRGAQTYPGEAGERWFG RRRRGRVVSRKKMSLKSERRGIHVDQSDLL
						CKKGCGYYGNPAWQGFCSKCWREEYHKAR
						QKQIQEDWELAERLQREEEEAFASSQSSQGA
						QSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
						VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE
1 1						FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE
			ľ			EQSECAQDFYHNVAERMQTRGKVPPERVEKI
						MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI
j						QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV
						KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI
						KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI
1 1						QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE
1					•	KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES
						WSPDACLGVKQMYKNLDLLSQLNERQERIM
]]						NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI
1007	0415		0077	1270	1610	KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
1067	2417	Α	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
1068	2418		8856	1620	1502	LRQAWATKQDPISKKK
1069	2419	A		1530	1583	PCRPGMECNSMISVHCNL
1070	2419	A A	8857 8866	1530 293	1583 1675	PCRPGMECNSMISVHCNL PVPOGGVPOGDVPOGGVPOGDVPOGGVPOGDVPOGGVPOGDVPOGD
1 10,0	2720	Λ.	0000	273	10/3	PYPQGGYPQGPYPQEGYPQGPYPQGGYPQGP VPOSPEPPNPYGOPOVEPGODPDSPOHGNYO
						YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
						LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
						WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL
		ì		ļ		VALSVLTASLSYMVGMIASFYNTEAVIMAVG
			j			ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
				l	ł	VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
				ļ		VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI
1	Í	Ì	ļ		ĺ	FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP
			Ì	ĺ		WHGSASCTSPLSCPQAQPREKDASLQPSCMY
			ĺ	ļ		TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
		- 1	ļ	-	1	HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
			[ļ	٠. ا	EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
						GDMRSGGLIPVLSPE
1071	2421	Α	8868	2	358	ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH
		J	J	ļ	ļ	DDKMGSNTFFKRNDCRYVMISCKADMAYDN
		1	j			VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII
						LNGEKLKVFPVRSGT*QGCSVWP
1072	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL
	j		j			GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI
				.		GGTGPSSDAGWGCMLRCGQMMLAQALICRH
ــــــــــــــــــــــــــــــــــــــ					l	LGRDWSWEKQKEQPKEYQRILQCFLDRKDC

SEQ ID NO: of	SEQ ID	Met	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	NO: of peptide	noa	in NO:	beginning nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
	 	 	 	sequence		CYSIHQMAQMGVGEGKSIGEWVLGPNTVAQ
						GV*KNLA/LFDEW/NSLGLVYVSM/DNPSGSIA
		}			}	RFPKKLCRVLPL\SADTAGLTGP
1073	2423	A	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL
		1				*QVCITA*IKESVIISGG*SSSPVCHTTFQPANL
		<u> </u>				RTSRYLPT*SIKSLGPDEPQEG
1074	2424	A	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW
						KEISFGDYICHTFQGDCWADRSPLHEAAAHG
						RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL
1075	2425	A	8896	1294	248	*GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	^	8670	1254	240	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
1						PWPSLLDKEREESLROKRLSERERIGELGAPE
	1					VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
						TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK
						KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK
						EKKKKHRSKKYKKKRSKKSRKESSDSSSKES
						QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS
			1			QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
						PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR MEAVRTAKREPESTVLMRREPLHPFNPRRET
						KERE .
1076	2426	Α	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
						*APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
						FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE
j · ;			}			VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR
						VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE
						LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK
1077	2427	Α	8901	352	3	RSQREHVQQQSQEHGKWPDLKGPR
10//	2421	A	8901	332	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW QYPALHRAGTEWQLSALHRAPRSTOPDKAC
			ĺ			RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
						\YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	Α	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA
				i		PALPFAATPGSRGQALCRGGRRRQHLHGPLH
						RP*QAAPALHAGCQLAPHPPT
1079	2429	Α	8912	121	376 -	NLIWKLCYTERRLVILDNYDLASE/YEANKYI
						CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI
1080	2420	Α	8020	201	1700	EYNKNGHLSFKYVKTFSMDEY
1000	2430	A .	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
		-		}		DOOTAALAMEPFHPMVNLDCSRDFRPFLCAL
		1				YAPICMEYGRVTLPCRRLCQRAYSECSKLME
1 1	-	-				MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA
						GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY
		l				SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS
		ſ	Ī			IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV
				1	•	WHMMVSLIFF\GFLLEDRVACNA\SIPAQYKA
	. 1	}		1		STVTQGSHNKACTMLFMILYFFTMAGSVWW
	Ì	1	l		ł	VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD
]]		j	l		ļ	VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR
		- 1	ŀ	l		VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL
	1		ŀ			VVIGCYFYEQAYRGIWETTWIQERC
1081	2431	A	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE
				1		CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG
	l	l	ľ	1		TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
1000	2422		9000	755	1070	LSIAIPAMVNNTAPPSQPNASTERPST
1082	2432	Α	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		Ì		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ı				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]]	peptide	ļ	/-possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
1						GPFSKKDQYDVKAPAMFNIRNTGK/TLVART
					i	QGTQIASDGLKGLLFEVSLADLQNDEVAFRK
1		ł			<u> </u>	FKLITEDVQDKNCLTNFYGMDLTCDKICSMV
						EKWSTMIEAHVDVKTTDGYFFHLFCVGFTKK
						HNNQILKTSYA*HQQS/RQIQKKMMEIMT*EV
1						QTNDLKEVVNKLIPDNIGKDTEKV/CPIYPLH
						DVFIRKVKMLENPGFER\MELRGGGSSS
1083	2433	A	8948	28	385	LTWPQPHIPSCPAMSEETLQSKLAAAKKKLP
1		Ì				WGAVQGSRAMSDLLLLLLDLTLLLLLMLLGF
1						AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
						GWLLT/ESCSISPKLCSIAVH*DNPAWF
1084	2434	A.	8950	156	318	HYTPINTDTIENSENNKCW*GY*E\VGLIHHW
1005	0.10.5	L			L	WGGKRVQPFWKRVWQKRTLNLRV
1085	2435	A	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK
						*TTYTSYDTAIPIS/GI/YPKRMSSKCHQETCAR
						MFILAPFTATIKGKQLTCPLVEERIDY\MWYS
1				!		HKYYIKVKRNL*VTITH\TWVNLNILMFEIILW
1004	2424		20.60	0.60	1005	YSHKYY
1086	2436	A	8962	868	1026	H*KILQVGRAQRAHXSRL*SQLLRRLRHESHL
1005						NPGARGCSEARLHRCTPAWTT
1087	2437	A	8985	58	330	LHVKHLGHFQLVFSEVICHCILMPVS*ELQRL
1 1						*ERSVCAFHVCIQTYVCLQVYACMCVYYICM
1000	2420		0000	204	404	FVYSVYGCGLCTCVCMDVYICVCVQEFL
1088	2438	A	8989	394	404	N*KWILHVNVRIQSIFF/IKRNQK/INSHELKLD
						KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA
1						KYTVKRIKIHPTDLEKMLRNHLSDKD*YS/GV
i i						YKDLSKLNRRKTE/S*/VKKWVKDLSRYFIKE
1089	2439	A	8991	60	329	VISMENKHKKIFSTS MALTPESPSSFPGLAATGSSVPEPPGGPNATL
1009	2437	Λ.	6991	00	329	NSSWDSPTEPSSLEDLEATGTGTLLSDMGVV
1 1						GVEDNAYTLEVNSRYMRAVGIM*IHL
1090	2440	Α	8996	2	351	SNITITLT*MKKYDNTFCW*GCGQIG/T/LIYC
1 1050	2110	: •	6770	-	331	WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL
1						GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR
1 1						LLIAALFIIVQYWKQSKDHYI
1091	2441	Α	8997	97	456	YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT
		**	3,,,		.50	LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG
						AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS
1						FDKNYKAPIGADFEMERFEVLGIPF
1092	2442	A	8999	548	811	SSFIKRHILIFEDDWHQTTCCHHPHHP\F*RCQ
		••	3,,,			FHIFYVSVONSISPSLSVSSSHPDRPDHEVHOH
]						RAAHHQHGQGPLGHGLVARVG
1093	2443	Ā	9002	3	2745	ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS
	2.15	••	1002	-		TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD
				1	 	AMCCLRYWYTPESWICGGOWREYFSALRDF
	İ					VKTFTAAGIKLIFFFDGMVEODKRDEWVKRR
						LKNNREISRIFHYIKSHKEOPGRNMFFIPSGLA
		[[VFTRFALKTLGOETLCSLOEADYEVASYGLO
			l			HNCLGILGEDTDYLIYDTCPYFSISELCLESLD
	1		[TVMLCREKLCESLGLCVADLPLLACLLGNDII
	ĺ	ľ		ĺ	ľ	PEGMFESFRYKCLSSYTSVKENFDKKGNIILA
				l		VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL
] 1		ļ		l		*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP
1 1		1				RVOTPNPGKKFPCVOMLNPGKKFPCVOALNP
J]						GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG
		ĺ	[[[PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP
			l			MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP
			İ			RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY
j 1	1	ļ	. 1	İ	ľ	TGPESRQEVLIRTDPESRQEIMCTGHESKQEV
		ــــــــــــــــــــــــــــــــــــــ				

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PICTDPISKQEDSMCTHAEINQKLPVATDFEFK LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS
1094	2444		9021			DTEILKVARTHHVQAESYLVYNIMSSGEIECS NTLEDELDQALPSQAFIYRPIRQRVYSLLLED CQDVTSTCLAVKEWFVYPGNPLRHPDLVRPL QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA CFNLSSSREELQAVESPFQALCCLLIYLFVQV DTLCLEDLHAFIAQALCLQGKSTSQLVNLQP DYINPRAVQLGSLLVRGLTTLVLVNSACGFP WKTSDFMPWNVFDGKLFHQKYLQSEKGYA VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV
		Α		97	834	AREACRAKTDFPGRRFRLWPSCCCRVIVGAE T*HMAEPVSPLKHFVLAKKAITAIFDQLLEFV TEGSHFVEATYKNPELDRIATEDDLVEMQGY KDKLSIGEVLSRRHMKVAFFGRTSSGKSSVI NAMLWDKVLPSGIGHITNCFLSVEGTDGDKA YLMTEGSDEKKSVKTVNQLAHALHMDKDLK AGCLVRVFWPKAKCALLRDDLVLVDGPGTD VTTELDSWIDKFCTKSSTREITNSGSDT
1095	2445	A	9022	1	537	LVLNSRVEDFVPPEGAGRTLPFALRPLAACW LLHRRARRSSALCPRPRSWGVSGGEGAGARE P*ITSSSCCLSAA/SHLSIQSPNMAGARRIRPQ LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV FAYGQT\GAGKTYTMGTGFD
1096	2446	Α	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL GQHSETPSLKKK\LAGYSGMCL*SQVLRRLRQ EDCLSPGGGNCRES*SCPYTPAWITERDPV
1097	2447	Α	9032	716	357	ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI LAHCNFRHAGFPPLSCLSLPNRWEYRRPPARP GKPFLVFLVETGFQC/G*DGLDLLTSRSACLG LPKCWDYRREPAASIIFQTTFFINSK
1098	2448	A	9038	230	652	KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK
1099	2449	A	9043	185	372	IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI
1100	2450	A	9045	763	584	RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF
1101	2451	A	9050	275	2	LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS
1102	2452	A	9053	449	1224	KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV
1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA AGMQI/H/CW\WCVNVGKFWEMS*YYLLKLSI ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion APFVLAVNC
		A		75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V KTDCGCGANSKGVVVVMKV\KTAQQKQTTS YMQIGTTKNSRAT
1105	2455	A	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR AWPCCPGWSAAWLTIVILAHYRRPGLERSCC LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL VLNS*TQGI
1106	2456	A	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT* AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	A	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA SAFPPAERSRGHRRASL*RARWSAAVPRRSA GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG QRPPPPSGDSLSPPGCCRY
1109	2459	Α	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL GAVAHSCNPSTLVGRGGRITRGQELR
1110	2460	A	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP AAAGDPASLDFAQCLGYYGYSKFGNNNNYM NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG VMPPAQLTTINQSQLSAQLGLNLGGASMPHT SPSPPASKSATPSPSSSINEEDADEANRAIGEK RAAPDSGKKPKTPKK
1113	2463	A	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH VGQAGHEPLTSGDPPASASQSAGITGVSHQA WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS LLKCWDY
1114	2464	A	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII YLYVQTVPQRV
1115	2465	A	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE WRLQHKGRGRGDLHLPDHHLSVPSSADHPA QQPSQFNGRNLYFLPLFR
1116	2466	A	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA GAGGP*GSPAGRACGAAGCRPRPPRPAASSA *NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	I Amino acid company (A-Alexina C. Com)
NO: of	NO: of	hod	ID NO:	beginning		Amino acid sequence (A=Alanine C=Cysteine,
NO: of	I .	iroa		nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1	peptide	1	in USSN		location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1		location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l.	09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
dence			914	amino acid		Q=Glutamine, R=Arginine, S=Serine,
.[l .	ĺ		residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan,
			1		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ľ				peptide		/=possible nucleotide deletion, \=possible
		 		sequence		nucleotide insertion
1					1.	CPARTSVQGGTWTC*APAGRPAGLGGWEAE
1117	2467		9141	380	939	RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS
1111/	2407	A	9141	380	939	KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP
]		j		•	,	TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED
		i			İ	*CCLMASEASWTIT\ELWVTLTVEGKSVP/CL
	Ì					NTEATHSTLPSFQGPVSLASITVVGIDGQASKP
	ĺ	[[ſ	LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG*
1118	2468	_	0164	471		GILTKLSAFLTIPRLQPHLIAALSPSS
1110	2408	A	9154	471	2	AAGQVVVEVTSHLYLCITSDAAGLRLLPPAËS
1		l				ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP
1	[[[[EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ
		i				LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE
1119	2460	 	0155		2107	TKEKTNKLNFIKKVKRQLAEWEKIFQIA
11119	2469	A	9155	2	3187	ACPRLARRRRRVRSLRRRRGWLRARWSRGQ
				,	1	NNMAARRITQETFDAVLQEKAKRYHMDASG
		1		,		EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS
1	l	1				DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR
1		l			:	SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG
1		1				HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ
1						ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK
ł	!	Ì	1			ESRDYDVDHPGEADSV/LRGGSQVQARGRAL
						NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL
				i		RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD
			•			VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR
1			ĺ			KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK
[FSQLFQTLFELETETCAKMLASFKCSLKPEHR
i l						DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG
						AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP
ĺ						LLMACNAYELSVKMKTLSNPLDLALALETTN
						SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ
1						DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS
						SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK
						APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS
						LSPKERTLLKEDPAYWFLSDENSLEYKYYKL
						KLAEMQRMSENLRGADQKPTSADCAVRAML
						YSRAVRNLKKKLLP\WQRRGLLRAQG\LRG\
}						WKARRA\TTGTQTLLFLRAPGLKHHGRQAPG
						LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL
				ł		EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT
				l		METAEKLARFVAQVGPEIEQFSIENSTDNPDL
						WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH
				l	ı	NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP
(l				ł	ĺ	PREAELESPEVMPEEEDEDDEDGGEEAPAPG
						GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ
ļ ļ					J	ASSGTCFPRKRISSKSLKVGMIPAPKRVCLIQE
1						PKGECPPVGTVASSTVLGWWAVRVRRDRWR
1106	0.156		0.00			HFNPKEFCAPLQNVSRHSCFPVV
1120	2470	A	9163	124	207	PPRACRPCPRACPCPPT*KCSQPVSWPC
1121	2471	Α	9166	272	523	PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK
		ļ		j	i	V/CSHITDSLKFIGKGWVGMVTHACNPGTLG
ا ــــــا						G*GGWIA*VREFETSLGNM
1122	2472	C	9170	442	236	MNRRRFLRPADCHSGMRGTENGACSEGESQI
		i				HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT
						FSKRQN*
1123	2473	A	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE
		ľ			ļ	AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV
			j	· .	1	SDAAGQGVAITGNQTFNNWNWPNAMIFAAT
		ł		1	1	VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC
]	l			LTWINALGKFFG

SEQ ID	CEOTO	Met	CEO	Predicted	Deadisted and	LAST CONTRACTOR OF CONTRACTOR
NO: of	SEQ ID NO: of	hod	SEQ ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	corresponding to last amino	
uence	uance		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uena			214	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ļ	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
į.	ļ			peptide	sequence	/=possible nucleotide deletion, \=possible
1		l		sequence		nucleotide insertion
1124	2474	A	9173	3	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
11124		^	7173	3	} 3/4	WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
Į.		ļ				TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA
1		l	1			IAHRGGRHDPPENTLGAIR/OGS**WSNRR
1125	2475	A	9179	704	188	ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL
125	1	^	7.,,	704	1 100	CSKPPKETGELENAESGGDGGRRGGKQDNV
1	Ì		İ		İ	AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA
	1					LPMGFFYLYFRDPGREITWKHFVQYYLARGL
J	j	ļ				VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE
		1			1	YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	Ā	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM
11121	2411	^	3103		321	EYMAESTDRSPGHILCCECGVPISPNPAQY\CV
			1		•	ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY
						LTFFSTIS
1128	2478	A	9186	183	847	
1126	2470	^	7160	103	047	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC
1	1		1			
1		ļ] .			L\LDGVPVALKKVQIFDLMDAKARADCIKEID LLKQLNHPNVIKYYASFIEDNELNIVLELADA
						GDLSRMIKHFKKOKRLIPERTVWKYFVOLCS
		i	ļ			ALEHMHSRRVMHRDIKPANVFITATGVVKLG
		ŀ				
1	ì		}			DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD NG
1129	2479	Ā	9190	1	370	
1129	2413	A	3130		370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
1			•		'	RATT/KIRVVATTTRARIEDMRHSATALTRPD
1						ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
1130	2700	A.	7174	131	40/	LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN
			ļ			PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
						DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE
1	2701	Λ.	3201	104	003	PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
ł .						CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE
						GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL
		•				AATGVYIFFQNKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIOLTTK
		•	/200	•		TQPVEATDDAFWDQFWADTATSVQDVFALV
				ŀ	·	PAAEIRAVREESPSNLATLCYKAVEKLVOGA
				l		ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD
]]	J			l		WRGFFWSTVPGAGRGGGGEEDDEHARPLAE
			[[SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH
1				ĺ		SLDSCEYTWEAGVGFAHSPOPNYIHDMNRME
				ļ		LLKLLLTCFSEAMYLPPAPESWQH/RTHWFSS
! !				ŀ		FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY
					ļ	NHLY .
1133	2483	Ā	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV
	2.05	**	7200		.700	AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
	ľ			ì	ł	HPOPSRGNPVGCLPTYKVVYKLLSWPLHSNS
						NVYFIV
1134	2484	Ā	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK
	2707	1	72.0	"	1200	RADGAAPAGEGEGVTLQGNITLLKGVAVIVV
	ĺ		l	ļ		AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC
		- 1				GVFSIVGALCYAELGTTISKSGGDYAYMLDV
		İ	1	Ì		YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL
]]	ļ		J	ļ		LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
j						YSVKAATRVQDAFAAAKLLALALIILLGFVQI
						GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG
	į					LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS RLFFVGSREGHLPSILSMIHPQLLTPVPSLVFT CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII GMIWLRHRKPELERPIKVNLALPVFFILACLF LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV WWKNKPKWAPPGHLSPRPSCVRSSCMVVPQ RDRLPPAYFCRPVVCVVTALDVG\SPESOEM
	2403		7210	70	410	DLVAFEDVAVNFTQEEWSLLDPSQKNLYREV MQETLRNLASIGEKWKDQNIEDQYKNPRNNL RSLLGERVDENTEENHCGETSSQIPDDTLNK
1136	2486	A	9223	3	983	RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG DLVFAKMKGYPHWPARIDDIADGAVKPPPN KYPIFFFGTHETAFLGPKDLFPYDKCKDKYGK PNKRKGFNEGLWEIQNNPHASYSAPPPVSSSD SEAPEANPADGSDADEDDEG'RGVMAVTAVT ATAASDRMESDSDSDKSSDNSGLKRKTPALK MSVSKRARKASSDLDQASVSPSEEENSESSSE SEKTSDQDFTPEKKAAVRAPRRGPLGGRKKK APSASDSDSKADSDGAKPEPVAMARSASSSSS SSSSSDSDVSVKKPPRGRKPAEKPLPKPPGRK PKPERPPSSSSSD
1137	2487	A	9229	21	239	LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA
1138	2488	A	9231	207	443	TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL EGIVWHETEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGKR ARSAAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGIIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGOPWRRAAAAGILPGREAAACLPSC/AS
						VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL
1140	2490	Α	9238	248	328	MAQGNNYGQTSNGVADESPNMLVYRKV
1141	2491	Ā	9242	2	535	FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVHCYSKSPSNKDAALLEAARAQ WMQEVSRNRCALLHSAAVQEYGYGN
_	2492	Α	7243	157	· .	HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ
1143	2493	A	9247	264	115	GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG ARDSTSIIRMGPEIPPP

SEQ ID NO: of	SEQ ID	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	NO: of peptide	noa	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	j	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciice		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence	ł	1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	}			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1	1	İ	l	peptide	sequence	/=possible nucleotide deletion, \=possible
	1	l		sequence		nucleotide insertion
1144	2494	A	9260	Sequence	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP
1144	2777	^	7200	1	401	SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE
1 !	ĺ	ĺ	1		ĺ	AMEESDRPCEISEIDDNPKISENPRRSPTHEKN
1	ľ		}		ļ	TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM
l					İ	ERRR
1145	2495	A	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF
1145	2493	^	1204	173	*''	PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL
		١.				WDTAGQERFISIT
1146	2496	A	9277	592	814	MFTYLEGREGIKSQPKMEPHSVT\RLECSGMI
1170	2430	Ι Δ	3211	392	014	SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA
						WLIFAFLVETGF
1147	2497	A	9279	1255	2	FRRGRRGEEKEEEEEEGWVNGMENSHPP
114/	2491	Λ.	7217	1233	4	HHHHOOPPPOPGPSGERRNHHWRSYKLMIDP
į						
i I						ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE DPRVVGIWTKNKE\LELSVPKFKIDEFYVDOV
ł						PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE
1						VEILYNPKTKKHLGIAKVVFATVRGAKDAVO
)	ļ]	
						HLHSTSVMGNIIHVELDTKGETRMRFYEL\LV
						TGRYTPQTLPVGELDAVSPIVNETLQLSDALK
[RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA
į l		i .				YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY
						NRRHEHHYVHNSPAVTAVAGATAAFRGSSD
1				•		LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP
1						PAQATPAPGFR
1148	2498	A	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY
1170	2470	Λ.	9302	1020		ADPONLTDVSIFLLLEVSGDPELOPVLAGLFL
						SMCLVTVLGNLLIILAISPDSHLHTPMYFFFSN
ł i						LSLPDV\GFTSTTVPK\MIVDI\QSRSRVISYAG
j i						CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI
						CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSOL
						HSWIVLOFTIIKNVEISNFVCDPSQLLKFACSD
i						SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL
						GNLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT
						STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF
1 1						GGMEERHAPECDGL
1149	2499	Α	9303	1	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE
				-		FRLVAADRSMGRYMLFGVINLICTGFLLMWC
J						SSTNSIALT\SYTYLTIFDLFSLMTCLISYWVTL
1						RKPSPVYSFGFERLEVLAVFASTVLAQLGALF
1		l l				ILKESAERFLEQPEIHTGRLLVGTFVALCFNLF
		J	. [TMLSIRNKPFAYVSEAASTSWLOEHVADLSR
						SLCGIIPGLSSIFLPRMNPFVLIDLAGAFALCIT
						YMLIEI
1150	2500	A	9308	797	693	DRSTSVTRAGVQWCSLGSLOPRTPGLLRSSCL
						SLP
1151	2501	A	9309	205	406	VAIKELPVLWKWSKPTR\TAKEPPOTOORAG
		•-				SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ
						FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPOVEV
		[7517	·••	207	QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR
[]	1	ł	ł	ľ	ľ	PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I
		i	!	ļ		PPTSA/FPKCWEYRSEPOCLPGCLSFSGILLDL
	.			ļ		GTNVSLRAA
1153	2503	A	9315	392	1	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP
1133	2503 [^	2010	374	,	
		ļ	İ			PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG
			.			GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR
					ļ	PGNS PGNS
	1					1 0110

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide		in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-)	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	·	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Í		ſ	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide	1	/=possible nucleotide deletion, \=possible
1154	2504	A	9321	sequence	433	nucleotide insertion
1134	2304	^	9321	331	433	MPCVQAQYGTPAPSPGPRDHSASDPLTPEFIK PT
1155	2505	A	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	Ā	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
					•	EAAQRHCRASVSILRMRRPGOGSSRPARVPL
						RGPDSHRLREPPPSPP
1157	2507	A	9327	152	292	YERRGRSQGGSHPAGAQPGGRAIGAGWQS
		L	<u> </u>			KEPLWEGLQRSGSPLPG
1158	2508	Α	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
						LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
}		1	i '			VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
}	1	1	Į,			SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/ RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR
1,	2009	1	2224	100	363	LMEAGLPQKQAERADELFEAGLVIYVKLDER
						VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	A	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
		ļ	Ì			KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
	1					DHTDQELREEIHKANVERVVHDVSQEATIEKI
1	l	ľ	1			RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
						EAELPIMSQLTEIETCVEC
1161	2511	A	9341	1	390	NSRVDDFVAPGLSEAGKLLGLEFPERQRLAA
		1				AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
		•				DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM
	1	}	1			QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT GELI
1162	2512	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
102	=====	**	/345	0.	037	TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
						VVEPISDEDWYLFCGDTVEILEGKDAGKOGK
	ł		i			VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
						GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
						FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
	}					WIDGPKDTSVEDALERTYVPCLKTLQEEVME
						AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE
1163	2513	A	9346	967	<i>(</i> 1 <i>(</i>	G DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
1103	داسا	^	3340	201	616	LPSSWAYRCASPHPDNFFVFLVESGFHHVGO
[1	AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
					1	FSSYQRNPPDLILNDTIMPNIK
1164	2514	A	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI
						HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL
			t l			FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV
						VAGASVGAGVWARNPRYRTEGEACVEFKA
]]			MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL
						VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
						SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
						SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS
{		1			Ì	YVSIFVPLWLSLLTLMATTFRRKGGNHWWF
}						AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA
			}			LPLONKDRGSWPASRGSPRLL
1165	2515	A	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP
				Ì	1	VPEGVRLADGPGHCKGRVEVKHQNQWYTV
	·				ı	CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
						TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
1122			1			LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	A	9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS
1167	2517	Α	9368	707	1007	GAISAHLAHCNLCLPGSSDSPASAFQVAS
110/	2517	<u>A</u>	2200	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid	to last amino acid residue of peptide	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						PRPLILYAPAPARPAGTAFIPHSHPPPPDLLRPT ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1168	2518	A	9375	511	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS QQSILAGLVVVATTGMIGSPLECLFGELGGRA DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL FNAMQASSEH
1169	2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN NSARRMEAMASGSNWLSGVNVVLVMAYWS LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH NAPKDLKEEIDILLSRVHNIKYEPVHLLADDDA
1170	2520	A	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF TNETWQARTGEPLPDHLVLLMWSLIVSLYPL GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS AAILFGFSRKAGSFEMIMLGRLASWGVNAGV SMNIQPIMLPGGESAPKELRGAVAMSSAIFTA LGIVMGQVVGLSTTAATGLRGL\AGELEELEE ERAACQGCRARRPWELFQHRALRRQVTSLV VLGSAMELCGNDSVYAYASSVFRKAGVPEA KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL WGGTPRSFALNQFILQKKKK
1171	2521	A	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG SKRLFFWMQEP
1172	2522	Α	9384	20	355	GWNGRSTEASPAAEAPHVPHKETKAAMGTQ CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL ASFNEVGNTALIVLESY
1173	2523	A	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI* KIFVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV IRPPISFSKINNGP
1174	2524	A	9397	77	374	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF LGML
1175	2525	A	9399	66	397	HESSRADRDKMDTRGSTYTDADPVNKSGGT AKMNKWSKGKVRDKLNNLVLFDTATYDKL CKEVPNYKLITLAVVSERLKIPGSLARAALHE LLSRGLI*LVIQHIAQVIY
1176	2526	A	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY GERGYAQNGDF*DAQLDDYSFSCYSHAQVN GAPNSLTRAYDDP*VKISGLECQKVGALVEV KCLNL
1177	2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1178	2528	A		142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179	2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA WWRAP

C050-	1 000 00	1.2.	1 000			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	испсе		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	}	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	į	l	peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1180	2530	Α	9422	176	375	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP
į.		ļ				SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK
						L
1181	2531	A	9436	2	274	PIAASLRMYNLQPYTEENLICTAFATMVETVP
1		1	1			IARTILDRLTGIPHGYCFVE*ADWATADKCVH
		l	<u></u> _			IYNGKPLPGATPLLSLQLHQLAHLGS
1182	2532	A	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL
	l		1			SMILK*MGAGDEKISAMGKARVDHRELYLGL
L						LYPTEDYKLTFRARH
1183	2533	A	9444	384	3	LKDFQPWALHDWPLFCCCTFLLFLVLECFTR
Ì						KGCSGWAPWLSLQCQHFGRPRWADHLRSGV
	1		1			RDQPGQYSKTTFLPKIQKLAGHSGAHL*S*LL
		ŀ				ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT
						ERG
1184	2534	A	9462	391	655	LSGFKSLMPKIPLQYTYVRVRTTWSFCLPLDG
1	l	1	ł i			RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV
			ļ			IHTCNPSTLGGRAGWIV*AQEFET
1185	2535	A	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA
			' ' '		500	WWWGWECWVRALKLSSGPAGPLACWVAK
						KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG
		i				WSAVVQFWLTAASNSCFSLLSSWDYRCA
1186	2536	Α	9468	275	452	HIPOLHTKTHYVPTRMVNKI*QIDNSKPWOR
	2000	1	7100	2,3	132	GG*TGILTHCW*ESKLVQPLWKIVWHYQ
1187	2537	A	9469	388	3	EVAPGPSQILPRRVTDGGDRPQFSLPGPRLPQ
1107	2007	^	7407	300]	SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP
						NPASPHPEAPQEPWDSASGSVGSFSLGRGAK
					Í	ASS*VPGKGRGPRQGSELLAETILELFLALAN
	}					S
1188	2538	A	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT
1100	2030	^	/4/1	124	371	GRLMANPEALKILSAITOPMVEEALAGLYRAC
		1				*FYLTNNLAGMKKGLCLGSTEQAHTIGI
1189	2539	A	9480	584	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET
1107	2239	^	7400	304	103	
j .	ļ	j			'	PSLLKIQKISWAWWRAPVVPATWEAEAEEW R
1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL
1190	2540	^	7403	403	00	
1]			PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR
						GASSCRRRRCNPVLAARKAGSPRSHSTRENC
1101	2541		0400		411	RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	A	9489	1	411	LADALCLSAAATGAVRPGARAQPSTRRRLSP
		•	'			SVRVCCRAAAASNLLYSSCLQRHSERASEEG
						ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI
•						MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ
1100	25.42		0400	200		KEEELTAVNVK
1192	2542	A	9497	389	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*
			, l	•		CEEDERKMAREFLAEFMSTYVMMNIHMIVE
			1			KDTYSDHEEINTS
1193	2543	A	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
			<u> </u>			FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194	2544	A	9512	58	433	PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL
j			,			LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA
[SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI
				i		RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ
1195	2545	A	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
						AGGGASLPVAAGSCAAAPHTEPGAPOHLLDC
						PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
						LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL*
1 1		1		i		PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP
						APGSGPCGATARPSRGGRAGGSRARRPIPPGP
				ļ		GTRRTPSGCQNPAASGG
			i	l		

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	(USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
ບະກວຍ			914	ng to first	acid residue	Q=Ghitamine, R=Arginine, S=Serine,
		1	İ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	,			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
1100-	2516	ļ	-	sequence		nucleotide insertion
1196	2546	A	9518	229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
}			1		j	AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA
1197	2547	A	9521	289	448	GYIGALFPMSGGWPGGQ
1197	2547	A	9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
1					ŀ	HAMRHLTGNTSMAIRFL*ADSRFQVQRARYE APNWKYKYGY*IPVDMLC
1198	2548	A	9524	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
1176	2340	^	7524	204	1	VKQEPNICHL*ETHFPFRDTYRLKEREQKKRK
1						SSYS
1199	2549	A	9546	1785	1943	GGRFKESKLTNAGWORNSFFIGPPKSIPWAA
1,	2317	'	/5.0	1705	1,743	V*QRGDGKNPGVTHLNRPVGTX
1200	2550	A	9548	186	1	VNAEKEF*KIOHYFMTKSONKLHIEHTYLKPI
1		٠٠.	/3.0	100	•	KAIYDKWTSDIMLNLQKL*AFFLRVIVRQI
1201	2551	A	9549	591	2	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
					_	GSCPSGE*LVSPGSEORKKYSNSNVIMHETSO
		l				YHVOHLATFIMDKSEAITSVDDAIRKLVOLSS
]]			KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
						PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK
1						PDVHFFHCDEVEAELVHEYMESALTDCRLGK
1						AMRP
1202	2552	Α	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
						LDCERPPQGPLPSLPELAKTSYSDLTGLATED
			1			*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
1 1						LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV
						SKPRATPPLFCSLHTF
1203	2553	A	9568	517	738	RRKFERKQKQ*RYREGKQYRQRDKMKEWG
						EKEKRREKGEREERKMRHRERKGESGQRD
1204	2554	A	9573	83	415	TMENWRVERLTEKER
1204	2334	Α	9313	83	415	EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
1						DGWDLSDAHVVCQKLGCGVAFNATVSAHFG EGSGPIWLDDLNCTGTESHLWQCPSRGWGO
						HDCRHKEDAGVICSEFTALR
1205	2555	Α	9577	64	424	ARGSCPTRPRTANGRMGETKDAPOMLVTFK
1 1	2000		3377	0.	727	DVAVTFFREEWRQLVLVHRTLYR*GMLETC
1					,	GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
l i						VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	Α	9584	38	476	TLGAVLFSEVSKESSTSHSGGOLGRONRHPKL
						SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
						NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
						GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
						YNPGLPPLRTWNGQKLLWL
1207	2557	A	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
		i				PDGCRNVLRPKYYRLCDKAESWGIALETVPT
						GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP
	i					THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
1200	2550		0505	122		FGILFSICFS
1208	2558	A	9597	122	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
1209	2660		0611	140		FADAWADAW
1209	2559	Α	9611	148	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
] [ĺ				GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ
						RIRDHDLLDKRKTVTALKAGEDRAILLGLAM MVCSIMM*FLLGITLLRSYMQSVWTRESQCT
		1			ŀ	LLNASITETFNC
1210	2560	A	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
	2000	<i>''</i>	2010	JUT .	-	DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
		ļ			:	KVQVKNNDLGLQATINNEANWIAHQDDFNW
1 1	ł	1		ľ		LLAELNTCQRQETADS***WSPKNSHVGKDS
L I			1	ļ		GELSAK WSI RAISH V GROS
1211	2561	A	9620	316	610	QKHPGGGQLGRSPQEDSRFHNKASSGVSRVR
				·		

SEQ ID	SEQ ID	Met	SEQ	Predicted	Deadisted and	I A-incode de la constante de
					Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	i			peptide		/=possible nucleotide deletion, \=possible
1	l				Ì	nucleotide insertion
}	<u> </u>	├	-	sequence		
İ		l		1		LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
-						GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE
Ł	l	L	1		ł	TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	A	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR
	1	1		l		TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP
J	İ	J		}		AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA
1	ŀ	1				DL*NTSFGVIR
1213	2563	A	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA
1213	2303	^	3024	2	330	
ł	i		i	i	1	TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ
l		}	1		l .	IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG
L	<u> </u>	L	L		<u> </u>	LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR
	1)		SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST
		١	1333	`	'~`	EENFGEKLHDIGFGNGFLDKT*KAOATKAKI
1		1	1 .			DK
1216	2566	 	0000	201	76	<u> </u>
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC
	}	1	ļ			RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI
1	ľ		ĺ			NHLPETERNLLEHGLMYIRLNAAFCSLVAHS
						LFGFILKAT
1217	2567	A	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV
1	1 200.	J	7025	2000	2432	EEHHLQPVQVLQTLLHSATAGTGCRRPARPP
			1			
ĺ			İ			PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL
						GALGGRGGRALGGSRWPPPLPGETLFSGCKH
						RRRRGSDAAPGEEAGT
1218	2568	Α	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA
İ	1					VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF
ļ	j	l	ļ			MLDFEGEDTFHGDMAKKETVWRLE*LARLD
	ì					NFEAQRALANIAADQAALEIMDMGSDYTLIP
1	l] .			
1010	2550		0.60		201	NVPPKVTVL
1219	2569	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA
İ		1	i			YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ
1	i	ł				LSQDPRNVWVFLATSGTLAGIMGMRFYHSG
	ł		1			KL
1220	2570	Α	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA
1	1 •			====		PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN
1	1	1			•	
1	1]			GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF
1		Ì]			PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI
1	l	}				SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL
L	<u> </u>	Ĺ				<u>K</u>
1221	2571	Α	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP
1						QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV
1	i				ļ	VOILTALMSLSMGITMMCMASNTYGSNPISV
1					j	
1					l	YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG
1000			L			SLGMNITSS
1222	2572	Α	9688	43	412	VAKMVKCCSÄIGCASRCLPNSKLKGLTFHVF
1					ł	PTDENIKRKWVLAMKRLDVNAAGIWEPKKG
			·			DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS
1.						PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	A	9696	308	564	RTSMGILYSEPICOAAYONDFGOVWRWVKE
1	ر ۽ صد	-	0,00	200	JUT	
}						DSSYANVQDGFNGDTPLICACRRGHVRIVSFL
<u> </u>						LKKECLCQPQKPERENLLALCCE
1224	2574	A	9700	3	632	DAWASGGELGSLFDHHVQRAVCDTRAKYRE
]						GRRPRAVKVYTINLESQYLLIQGVPAVGVMK
1				ļ		ELVERFALYGAIEOYNALDEYPAEDFTEVYLI
1			}	ļ		KFMNLQSARTAKRKMDEQSFFGGLLHVCYA
1			ļ į	}		
1				İ	1	PEFETVEETRKKLQMRKAYVVKTTENKDHY
1 1				ľ	i	VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA
L						ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Į		peptide	ĺ	/=possible nucleotide deletion, \=possible
		L		sequence		nucleotide insertion
1225	2575	Α	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
1226	2576	A	9713	82	492	TDDVHINDISLQDYIAGVRLILL
1220	2376	Α	9/13	62	492	QGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG
		l				ASVANKDIICYNLQAVGQIFYISSFLYTVNYI
ľ						WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ
1		ľ	i i		} ·	MAFVFSSLI
1227	2577	A	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG
		1				LAAGKMNISIDLDTNYAELVLNVGRVTLGEN
j					j	NRKKMKDCQLRKQQNENVSRAVCALLNSGG
İ						GVIKAEVENKGYSYKKDGIGLDLENSFSNML
1000	0.550		0000	250		PFVPNFLDFMQNGNYF
1228	2578	Α	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN
1229	2579	Α	9725	121	902	TLVLQKSDVEAVF LFAMSGFENLNTDFYOTSYSIDDOSOOSYDY
1227	2319	^	الكروا	121	302	GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP
		ļ]			QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP
						PLLEELGINFDHIWQKTLTVLHPLKVADGSIM
						NETDLAGPMVFCLAFGATLLLAGKIOFGYVY
			İ			GISAIGCLGMFCLLNLMSMTGVSFGCVASVL
	!					GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG
					:	WCSFSASKIFISALAMEGQQLLVAYPCALLYG
1000	440				· ·	VFALISVF
1230	2580	A	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG
1 1						HFSPERPFMDYFDGVLMFVDISGKCKRDVCL MWMSNRLAWEFTCRA
1231	2581	A	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP
1	2501	••		J.	1100	ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA
						LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS
1 1				•*		LRCGWSPAEELNYTVPGPGPAGEASPRQCRR
·						YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
j ,						RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ
		!				SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI
[NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG
						WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL LVLAGVAYALPHWRWLQFTVALPNFFFLLY
1						YWCIPESPRWLISQNKNAEAMRIIKHIAKKNG
						KSLPASL
1232	2582	Α	9753	164	517	PGPGMQGPPPTTPTSWSLPPWRAYVAAAVLC
	,			1		YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG
						LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS
				- · · · · · · · · · · · · · · · · · · ·		FGILLWSGAGLSSSFISPRYSWLF
1233	2583	Α	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT
]				İ		IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM
						KRAYKSYVRALPLLKKMGINSILLRKSIGALE
	1			ĺ		VACGIVMTLVPGRPKDVANFPLLLLVLAVLF FHOLV
1234	2584		9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP
~~		^	7,03	··	750	CSRCGYGVYPAEKISCIDQIWHKACFHCEVC
	ļ)	J	KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS
				i		VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV
-					ļ	FHWD
1235	2585	Α	9767	52 .	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA
1						SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI
						SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP
1	ļ		Į		ļ	ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP
]	-		1			CARPGHSARNNTDKNLPHTAILVTSNTYTTI
1236	2586	A	9770	352	600	KINFQAGRSGSCL FRGEALTVRFLTKRFIGEYASNFESIYKKHLC
1230	000 ع	^	2110	334	608	FROGALI VKFLIKKFIGE I ASNPESI I KKHLC

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			ľ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
			<u> </u>	sequence		nucleotide insertion
						LERKQLNLEIYDPCSQTQKAKFSLTSELHWA
						DGFVIVYDISDRSSFAFAKALI
1237	2587	Α	9793	266	515	NILAIIYFPFPRLFLLRDSQSNPKAFALTLCHH
						QKIKNFQILPVSIDALTPPLVVCFLVSFLTHFS
						RYKPTRPVCITQFQGCS
1238	2588	Α	9802	537	967	ELGAGRSDREAMEAAVKEEISVEDEAVDKNI
		İ				FRDCNKIAFYRRQKQWLSKKSTYRALLDSVT
1		Ì				TDEDSTRFQIINEASKVPLLAEIYGIEGNIFRLK
1		İ				INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT
				·		GSLILADGKGDLKC
1239	2589	Α	9805	105	540	VPGDPAMVRAGAVGAHLPASGLDIFGDLKK
1						MNKRQLYYQVLNFAMIVSSALMIWKGLIVLT
1		İ				GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED
1		!				PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG
						DIKFLTKGDNNEGDDRGSYK
1240	2590	Α	9819	3	305	TDGRDPLPCAARRRGGGGECCGAGWVAEWS
						PQPLDPAMLLWMQGFVLEAVACQDNDDYLR
		l				YGILFEDLDCNGDGVVDIIELQEGLRNWSSAF
						DPNSEEHG
1241	2591	Α	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
						LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR
]]						GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE
						SSHSVVEFLFKRTKTPSPFHPAVRENRN
1242	2592	Α	9843	3	589	TISCGPATEPPASLLSSASSDDFCKEKTEDRYS
						LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
						DGSVKCTHQPCLIKWISERGCWSCELCYYKY
1 1						HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
1						LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
1						MYGFMDVMIVAVDSEDMVQAAKEVGKRWS
1243	2593	A	9846	198	411	DIPP WRISHHAGKMPVMKGLLAPONTFLDTIATRF
1243	2393	Α	9040	190	411	DGTHSNFILANAQVAKGFPIVYCSDGFCELAG
			٠.			FARTEVMO
1244	2594	A	9848	116	650	PICGFLYLCSAMASESSPLLAYRLLGEEGVAL
1244	2374	^	7040	110	050	PANGAGGPGGASARKLSTFLGVVVPTVLSMF
i i			ł			SIVVFLRIGFVVGHAGLLOALAMLLVAYFILA
						LTVLSVCAIATNGAVQGGGAYCILQHRWTG
1 1						VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
			ļ			NVCGCAVSLLGLVESVLDVFGA
1245	2595	Α	9849	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE
		'	7047	5,5	1020	AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM
		l				SYIGPKRTAVVRGIMHREAFNIIGRRIVOVAO
[[•	AMSLTEDVLAAALADHLPEDKWSAEKRRPL
						KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR
						YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
		1				RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
ļ ļ	J	J	}	ļ	ļ	SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT
				l	l	NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
				l	!	EVDMVRVMEVFLAOLRLLFGIAOPOLPPKCL
		-				LSGPTSEGLMTWELDRLLWARSVENLATATT
[]	1			1		TLTSLA
1246	2596	A	9850	114	464	PPOLGAORVREPRHPDVRAPLRVTSPGLRSRS
						ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM
	Ì					QDGLSPCFFFTLVPSTRMALGTLALVLALPCK
						RRERPAGADSLSWGAGPRISSYV
1247	2597	A	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
		l	1			HQFPTDTIQRSKWIRAVNRVDPRSKKJWIPGP
•			l			GAILCSKHFQESDFESYGIRRKLKKGAVPSVS
<u> </u>	ĺ		. 1			LYKVFKYSSRCTS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence 58	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RVDDFVYSKGGKDAGGADVSLACRROSIPEE
	2376	A		36		FRGITVVELIKKEGSTLGLTISGGTDKDGKPR VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR LRHDEIITLLKNVGERVVLEVEYELPPPGGCP WT
1249	2599	A	9856	2		LPPPRPSRHRRGRAGTRASAAAAAGPTVSAV RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA SGSGVAAGPAARHAPRRCADAGEAVGASC GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAMRSTTLLALLALVLLYLVSGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFFSGTIITTIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFCIFYALVGIPLFGILLAGVGD RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVTLTTVGFGDYVA
1250	2600	A .	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF EWVYTDQPHTQRRKEILAKYPAIKALMRPDP RLKWAVLVLVLVQMLACWLVRGLAWRWLL FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV OLA
1251	2601	Α	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS
1252	2602	A	9879	6	376	KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	A	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC
1254	2604	A	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL
1255	2605	A	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT
1256	2606	A	9902	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE
1257	2607	Α	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence.	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH QRRGPSCGASGDPQCVGSPHPQRARPLLARP GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR PQGGHHPPLPTPGGRPCFAVSEGSGSALLLS YLGECGSSSYVTGAACISPVLRCREWFEAGLP WPYERGFLLHQKIALSRYATALEDTVDTSRL FRSRSLREFEEALFCHTKSFPISWDAYWDRND PLRDVDEAAVPVLCICSADDPVCGPPDHTLTT ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS HEVILESFRALTEFFRTEERIKGLSRHRASFLG GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL MAAAAGAAAAPGSREPQDRPECGAGHPGPR YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR ERPAARSGPEMRVRYPVVAAVLAPYLALSQD PMYKSSASGQGASGSYNHVREEMLIKAGGA MSRRVVRQSKFRHVFGQAAKADQAYEDIRV SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL PLAK
1259	2609	A	9919	693	935	GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL PPRPGRSHRKRKLVSTK
1260	2610	A	9921	455	1082	QRSCLCSAIEKDGGDVKALYRRSQALEKLGR LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE KKQKASQNLVVLAREDAGAEKIFRSNGVQLL QRLLDMGETDLMLAALRTLVGICSEHQSRTV ATLSILGTRRVVSILGVESQAVSLAACHLLQV MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	1	438	GFRGAEAPGAAQAPKKKKPRPTEGGPGAGSG RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF PTRVDHNGALLAFSPPFPQRQRRGTGATAES RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA PPRLPFCLQELQGRHALHTFSLERTCSYQDFL WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGFPYSLGEAAWSQL
1264	2614	A	9941	61	277	ESIGLTALGPRRRPWEHRWSDPITLKMK.GWG WLALLLGALLGTAWARRSQDLHCGACKAVR RRVRQFNIYDY
1265	2615	A	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLTILLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN
1266	2616	A	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	A	10005	2	209	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP SQDELEHSLGESAAQGAAGVVLWVSWENTR

nucleotide sequence uence unice contespondis of the contespondis o	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Desire	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Sequence 09/496 corresponding on the first of the side and side of residue of peptide sequence continue and incident of peptide sequence continu							
1269							
mino acid residue of peptide sequence s	1 -	June					
Peptide Sequence	20.00]	714			
Peptide	}	1					
	Ì	1					
1269 2619 A 10010 245 688 FÖMLKNIKHSSKKDNLAVNAVALODBILHD LOLRNLSVADHSKTOVCKKENSLKRITKAM LOTALKKTTOCPKLEDISKESVLUDEPPLIT. AOKLGLIGPPPPLSSDEWIKVKQRSLLQGDS VOPCPICKEFELROVPSING	1	1	ł	1		ļ	
1270		T			<u> </u>		TKVSLGLA
LQLRNLSVADHSKTQYQKKENSLKRDTAS	1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD
AQKLGLIGPPPPLSDEWEK/KQRSILQQDS	1	l	ł	1			
1270 2620 A 10011 2 588	l l						
1270			1				
PPEDVRGPSTHRTOY-VISPYDRPGGWNPEPCII SGNOLLM.DEDEHIPLIRDRRSSESSRNKLI.R RTVSVPVEGRPHGEHEYHLGRSRSKSVPGGK QYSMEGAPAAPFRISQGHJSRRLKSSIKTISS QPKLDRTSSFRQHJERYRSADHDRYRGWSMW DEIDV 1271 2621		<u> </u>					
SGNQLIMLDEDEIHPLIRDRRSESSRNLLR RIVSPYPGGRPHGEHVYHLGRSRRSVPGGK QYSMEGAPAAPPRE9QGFISRRLKSSKRTKS QPKLDRTSSFRQLPRFRSADHDRYRKGWSMW DEIDV 1271 2621 A 10013 209 363 LPAPPRISPRISFGQFFGGNDNYLTITGPSHP FISGAEVSQSCRRRGGRA LDEIDV 1272 2622 A 10014 7 388 SAVTISWKWRSVMGJOTSPALLASLGAGLVT LLGLAVGSYLVARSRRPQVTLDPNEKDLIS LDKTLSARSPCKHIYLSTRIDGSLSRPYTPVT SDEDQGYVDIDIKVYLKGVHZDGSLSRPYTPVT SDEDQGYVDIDIKVYLKGVHZDGSLSRPYTPVT DEDQGYVDIDIKVYLKGVHZDGSLSRPYTPVT DEDGGVFORDIKVYLKGVHZDGSLSRPYTPVT DEDGGVFORDIKVYLKGVHZDGSLSRPYTPVT DEDGGVFORDIKVYLKGVHZDGSLSRPYTPVT DEDGGVFORDIKVYLKGVHZDGSLSRPYDGSV SAPRAASGPSGSAPAVAAAAQPGSTPALS AQAAREPAAFWGPLARDILVDTFYHTVW DCDFSTGKIGWFLGGQLNVSVNCLDQHVKKS PESVALIWERDPEFGTEVRITYLGTFCRAMCAA CARIGAVHTVEAGFSAESLAGNDAKCKVV TIFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQGEMAKEDP VCAPESMGSEDMLPMLYTSGSTGMFKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYQFDLCNGATSVLTESTFVYPNA GRYWETVERLKNDGYTQAFTAVRLLLKYGD AWVKKYDRSSIRTLGSVGGFTCAWBUH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPGGTFRSPSSHVLTMSAPDEGRRDFPPKPG KTLGSFFOSLPGFSSARNLVANABSSARAPA ADPTGAPAAEAAQPQAQNAAHPSTELWFTE KELQPSEKMYSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVQGGLDTTKSALTGTKEV VSSGYTGAMDMAKGAVQGGLDTSKAVLTG TKDTVTSGLTGAVNVAARGTVQTGTGFTS KAVLTGTKDTVTCSGVTGAMNVAKGTTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTTQT ANANKGAMGTGLTTTSVLTGTKDTVCSGVTGAMNVAKGTTQT GVDTSKTVLTGTKDTVTCSGVTGAMNVAKGTTQT ANANKGAMGTGLTTTSVLTGTKDTVCSGVTGAMNVAKGTTQTGT TKDTVTTGTKDTVTTGTCTTTTCSGVTGAMTVAKGTVTGTG TKDTVTCTGTTTTTCTGTTTTCTTTTTCTTTTTCTTTTTCTTTTTT	1270	2620	Α	10011	2.	588	
RTYSPYPEGRHGEHEYH_GRSRRSYGGK QYSMEGAPAPPRPGQGTSRRLSSSRRTKS QPKLDRTSSFRQILPRFRSADHDRYRGWSMW DEIDV	1			<u> </u>			
1271 2621 A 10013 209 363	1	}	1				,
1271 2621 A 10013 209 363]]			
DEIDV			ł	ļ			
1271 2621 A 10013 209 363							
1272 2622 A 10014 7 388 SAVTISWKWRSVMGIQTSPALLASLGAGLVT LLGLAVGSYLVRSRRPQVTLLDFNEKDLLR LDKTLSARSPCKHYLSTRIDGSISRPYTPVT SDEDQGYVDDIKVYLKGVHPTPEGGKMSH LDKTLSARSPCKHYLSTRIDGSISRPYTPVT SDEDQGYVDDIKVYLKGVHPTPEGGKMSH AWARTLGRGVGRLLGSIRGISGOPARPPCGV SAPRRAASGPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDTLVWDTPYHTVW DCDFSTGKIGWFLGGQLNISVNCLDQHYRKS PESVALIWERDEPGTSEVRITYRELLETTCRLA NITLKHGVHRGDRVAITWPVSPLAVAAMLA CARIGAVHTVIPAGPSASLAGRIDDAKCKVV ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEGMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCHGEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCHGEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCHGVANVLFSSTFVYPNA GRYWETVERLKINOFYGAFTAVLILKYGD AWVKKYDRSKETLUTSVVGPINCEAWEWLH RVVGDSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWWQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTWXQT VVGPSRCTLVDTTKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAMNVA VVGRGLDTTKSVLTGTKDTVCSG VTGAMNLAGGTVQTGVDTTKSV VGTKDTMSTGLTGAMNVAKGTVQTGTVTTKVVLTG VKDTKDTMSTGLTGAMNVAKGTVQTGM DTTKTVLTGTKDTTVGLAANVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAMNVAKGTVQTGM DTTKTVLTGTKDTTVGLAANVAKGAVQTGT UKTTVLTGTKDTTVTGLMAGANVAKGAVQTGT UKTTVLTGTKDTTVTGLMAGANVAKGAVQTGT UKTTVLTGTKDTTVGLMAGANVAKGTVQTGM DTTKTVLTGTKDTTVGLMAGANVAKGAVQTGT UKTTVLTGTKDTTVGLMAGANVAKGTVQTGM UTTKTVLTGTKDTTVGLMAGANVAKGTVQTGM UTTK	1271	2621	_	10013	200	263	
1272 2622 A 10014 7 388 SAVTISWKWRSYMGIQTSPALLASI.GAGI.VT LLGLAVGSYLVRRSRRPQVTILDPNEKDLR LIBKTI.SARSPCKHIYLSTRIDGSLSIRYTPVY SDEDQGYVDILDPNEKDLR LIBKTI.SARSPCKHIYLSTRIDGSLSIRYTPVY SDEDQGYVDIJKVYLKGVHPITPEGGKMSH MAARTI.GRÜGGRILGSLEGGPARPPCGV SAPRRAASGPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDTLVWDTPYHTVW DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS PESVALIWERDEPGTEVRITYRELIETTCRLA NTILKRHGVHGGDGVQLNVSVNCLDQHVRKS PESVALIWERDEPGTEVRITYRELIETTCRLA NTILKRHGVHGGDRVALKKCVVV IIFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMFKGIVHT QAGYLLYAAATHKL VFDHQPGDIFGCVADIG WITCHSVVVYQPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPNCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTFRSPASHVLTMSAPDEGRRDPFKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAFFEQTAFWTE KELQPSEKMYSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVQGGLDTTKSALTGTKEV VSSQVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVCGGVTGAMNVAK GTTQTGVTSTS KAVLTGTKDAVSGVTGAMNVAKGTTQTGVETTS KAVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAK GTIQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT QVDTTKTVL GTKDTVCSGVTGAMNVAKGT QVDTTKTVL GTKDTVCSGVTGAMNVAKGT QVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGMT AKTVLTGTKDTTVTGAMVAKGAVQTGDTTKTVLTGTKDTTVTGAMVAKGAVQTGGM DTTKTVLTGTKDTTVTGLAMAVNAKGAVQTGGM DTTKTVLTGTKDTTTGLAMAVNAKGAVQTGGM DTTKTVLTGTKDTTTGLAMAVNAKGAVQTGGM DTTKTVLTGTKDTTTGLAMAVNAKGAVQTGGM DTTKTVLTGTKDTTTGLAMAVNAKGAVQTGGM DTTKTVLTTGTKDTTTGLAMAVNAKGAVQTGGM DTTKTVLTTGTKDTTTGLAMAVNAKGAVQTGGM DTTKTVLTTGTKDTTAGLAMAVAKGAVQTGGM DTTKTVLTTGTKDTTTGLAMAVNAKGAVQTGGM DTTKTVLTTGT	1271	2021	^	10013	209	303	
1273 2623 A 10016 1 1339 MAARTLGRÖVGRILGSLRGILSGOPARPPCGV SAPRANASGPSGSAPAVAAAAAQPGSVPALS AQAAREPAAFWÖPLARDILVWDTPYBIGKMSH DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS PESVALIWERDEPGTEVRITYRELLETTCRLA NTILKHGVHRGDRVAITWPVSPLAVAAMLA CARIGAVHTUFLAGFSAELAGRINDAKCKVV ITFNQGLRGGRVVELKKINDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITCHSTVVYOPLCHGNSTVLFESTEVYPNA GRYWETVERLKINGFYGAPTAVRILLKYGD AWKKKVDRSSNRTLGSGEPNICEAWEHH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPGGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFPGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPVTE KELQPSEKMVSGAKDLVCSKMSRAKDAVS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTKSVLTGT KAVLTGTKDTVTTGVMGAVNLAKGTTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKDTTVCSGVTGAMNVAKGT INTROTRUSTGLTGAMNVAKGTVQTGTVDTTKTVLTG INTROTRUSTGLTGAMNVAKGTVQTGTM DTKTVLTGTKDTTVTGLAMAVNAKGAVQTGD ITKTVLTGTKDTTTTGLAMAVNAKGAVQTGD ITKTVLTGTKDTTTTGLAMAVNAKGAVQTGD DTKTVLTGTKDTTTGLAMAVNAKGAVQTGD DTKTVLTGTKDTTTGLAMAVNAKGAVQTGD ITKTVLTGTKDTTTGLAMAVNAKGAVQTGD DTKTVLTGTKDTTTGLAMAVNAKGAVQTGD DTKTVLTGTKDTTTGLAMAVNAKGAVQTGD DTTKTVLTGTKDTTTGLAMAVNAKGAVQTGD DTTKTVLTGTKDTTTGLAMAVNAKGAVQTGD DTTKTVLTGTKDTTTGLAMAVNAKGAVQTGDM	1272	2622	A	10014	7	388	
1273 2623 A 10016 1 1339 MARTIGRÖYGRILGSLRGISGOPARPPCÖV SAPRRAASOPGSAPAVAAAAAQOPGSYPALS AQAARPAAFUGLARDILVWDTYPHTVW DCDFSTGKIGWFLGGOLNVSVNCLDQHYRKS PESVALIWERDEPGTEVRITYRELLETTCRLA NTLKRHGVHRGDRVATYMPVSPLAVAAMLA CARIGAVHTVIFAGFSAESLAGRINDAKCKVV ITIPNQCLRGGRVVGLKKIVDEAVKHCPTVOH VLVAHRTDINKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYYLLYAALTIHLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYFNA GRYWETVERLKINGYGAPTAVELLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTFRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFOSLPGTSSARNLVANAHSSARARPA ADPTGAPAAEAAQAQVAAHPEDTAPVTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGYTGAMDMAKGAVQGGLDTTRSALTGTKEV VSSGYTGAMDMAKGAVQGAVLAKGTVQTGVETS KAVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGTIQTG GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQTG TIGTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGTIQTG GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQTG TKDTVCSGVTSANNVAKGAVQGGLDTTKSVLTGTKD ANVAKGAMQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGANVAKGAVQTGVDTA KTVLTGTKDTTVTGLAGAVNVAKGAVQTGVDTA KTVLTGTKDTTVTGLAGAVNVAKGAVQTGVDTA KTVLTGTKDTTVTGLAGAVNVAKGAVQTGVDTA KTVLTGTKDTTVTGLAGAVNVAKGAVQTGVDTA KTVLTGTKDTTVTGLAGAVNVAKGAVQTG GLKTTQNIATGTKNTTGNGTVSTAVNVAKGAVQTG GLKTTQNIATGTKNTTTGNDTAVTTGLAGAVNVAKGAVQTG GLKTTQNIATGTKNTTTGNDTAVTTGAMAGAVNAKGA QTGVDTAKTVLTGTKDTVTTGLMAGAVNAKGAVQTG GLKTTQNIATKTVTTGTKDTAVTTGAMAGAVNAKGAVQTG	1) · · ·] ****]	•	300	
SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH 1273 2623 A 10016 1 1339 MARTLGRGVGRLLGSLRGLSGQPARPPCGV SAPRRAASCPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDILWDTPYHTVW DCDFSTGKIGWFLGGQLNVSVNCLDQHYKS ADAARACPGSTGKIGWFLGGQLNVSVNCLDQHYKS CARIGAVHTUFRGFSAESLAGRINDAKCKVV ITNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVTDHQPGDIFGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT RVVGDSRCTLVDTWWQT RVFQSSCRTLYDTWWQT APPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTKSAVLTG TKDTVSTGLTGAVNVAKGTIVQGTVETS KAVLTGTKDAVSTGLTGAVNVAKGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQTG AVSTGLTGAVNVAKGTQTGMT AANVAKGAMQTGLDTTKSVLTGTKD AANGACAQQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGAVQGGLDTTKSVLTGTKD KTVLTGTKDTVTGVGLVGAVANAKGAVQGGLDTTKSVLTGTKD KTVLTGTKDTVTGVGLVGAVANAKGAVQGGLDTTKSVLTGTKD KTVLTGTKDTTVTGLVGAVANAKGAVQTGVDTA KTVLTGTKDTTVTGLVGAVANAKGAVQTGVDTA KTVLTGTKDTTVTGLVGAVANAKGAVQTGVDTA KTVLTGTKDTVTGLGVANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVANAKGAVQTGVDTA KTVLTGTKDTTVLTGVANAKGAVQTGVDTA KTVLTGTKDTTVLTGVANAKGAVQTGVDTA KTVLTGTKDTTVLTGVANAKGAVQTGVDTA KTVLTGTKDTTVLTGVANAKGAVQTGVDTA KTVLTGTKDTTVLTGVANAKGAVQTGVDTA KTVLTGTKDTTVLTGVANAKGAVQTGVANAKGAVQTGVDTA KTVLTGTKDTTVLTGLANAKGAVQTGVANAKGAVQTGVANAKGAVQTGVATA GUTANAKTVLTGTKDTVTTGLMGAVANAKGAA GTVATAKTVLTGTKDTVTTGLMGAVANAKGAA GTVATAKTVLTGTKDTVTTGLMGAVANAKGAA GTVATAKTVLTGTKDTVTTGLMGAVANAKGAA GTVATAKTVLTGTKDTVTTGLMGAVANAKGAA							
1273 2623 A 10016 1 1339 MAARTLGRGVGRLIGSLGGQPARPPCGV SAPRRAASGPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDTLWDTPYHTVW DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS PESVALIWERDEPGTVRITYRELLETTCRLA NTLKRHGVHRGDRVATVMPVSPLAVAAMLA CARIGAVHVIFAGFSAESLAGRINDAKCKV ITRNQGLRGGRVVSPLEKKIVDEAVKHCPTVQH VLVAHRTDINKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITCHSYVVYGFLCNGATSVLFESTPVYFNA GRYWETVERLKINGYGAFTVAELLKYGD AWVKKYDRSSLRTLGSVGFPNCEAWEWLH RVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKFKG KTLGSFFOSLPGTSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAVAAHPEGTAPWTE KELQPSEKMVSOAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMMMAKGAVQGGLDTSKAVLTG TKDTVTTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQTGTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGTIQTGTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGTIQTGTGVDTTKTVLTGTKDTVCSGVTGAMNVAKGTIQTGTGTDTTTGTMATGTKDTTVCSGTGAMNVAKGTIQTGTGTTTKDTTVCSGVTGAMNVAKGTIQTGTGTTTKDTTVCSGVTGAMNVAKGTIQTGTGTTTKDTTVCSGVTGAMNVAKGTIQTGTGTAMNVAKGTQTGMDTKVLTGTKDTVCSGVTGAMNVAKGTIQTGTGTAMNVAKGTAQTGTAMAACAAVGAAMGAAAAAAAAAAAAAAAAAAAAAAAAAAA			l				
SAPRRAASGPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDTLVWDTPYHTVW DCDFSTGKIGWFLGGQLNVSYNCLDQHVRKS PESVALIWERDEPGTEVRITYRELLETTCRLA NTLKRHOVHRODRVANYMPVSPLAVAAMLA CARIGAVHTVIFAGFSAESLAGRINDAKCKVV ITFNQGLRGGRVVELKKINDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYOPLCNGATSVLFESTPYYFNA GRYWETVERLKINOFYGAPTAVRLLLKYGD AWVKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFGSLPGPSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKNVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGYMGAVNLARGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVAKGTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTTVCSGVTGAVN LAKEAIQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSVLTGTKD TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAGAVNVAKGAVQTGMDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTVTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTVTTGLMGAVNVAKGAVQT GLKTTQNIATGTKDTTVGLMGAVNVAK	1273	2623	A	10016	1	1339	
AQAAREPAAFWOPLARDTI.VWDTPYHTVW DCDFSTGKIGWFLGQLNVSVNCLDQHVRKS PESVALIWERDEPGTEVRITYRELLETTCRLA NTI.KRHGVHRGDR.VAIYMPVSPLAVAMLA CARIGAVHTVIFAGFSAESLAGRINDAKCKVV ITFNQGLRGGRVYELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPEMGSEDMIM.YTIGSGTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYOPLCNGATSVLFESTPYYPNA GRYWETVERLKNOPTGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTRSPSASHVLTMSAPDEGRRDPPKPKG KTLGSFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTKSVLTGTKEV VSSGVTGAMDMAKGAVQGGLDTTKSVLTGTKEV VSSGVTGAMDMAKGAVQGGLDTKSVVLTG TKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT QTOTSKTVLTGTKDTVCSGVTGAMNVAKGT QTOTSKTVLTGTKDTVCSGVTGAMNVAKGT QTOTSKTVLTGTKDTVCSGVTGAMNVAKGT QTOTSKTVLTGTKDTVCSGVTGAMNVAKGT QTGTDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGTDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGTSTTVLTGTKDTVCSGVTGAMNVAKGT QTGTDTTKTVLTGTKDTVCSGVTGAMNVAKGT QTGTSTTVLTGTKDTVCSGVTGAMNVAKGT QTGTSTTVLTGTKDTVCSGVTGAMNVAKGT QTGAMNLARGTIQTGVDTTKTVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTK VLTGTKDTVTTGLVGAVNVAKGTVQTGVDTTK VLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTVTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTTTGLAGAVNVAKGAVQT GLKTTQNIATGTKDTTTGLAGAVNVAKGAVQT GLKTTQNIATGTKDTTTGLAGAVNVAKGAVQT GLKTTQNIATGTKDTTTGLAGAVNVAKGAVQT GLKTTQNIATGTKDTTTGLAGAVNVAKGAVQT GLKTTQNIATGTKDTTTGLAGAVNVAKGAVQT GLKTTQNIATGTKDTTTGLAGAVNVAK							
PESVALIWERDEPGTEVRITYRELLETTCRLA NTI.KRHGVHRGDRVAIYMPVSPLAVAMALA CARIGAVHITVIFAGFSAESI.AGRINDAKCKVV ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVPJCLNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTILGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQTGVETS KAVLTGTKADVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT AANVAKGAMQGLDTTKSVLTGTKD AANVAKGAMQGLDTTKSVLTGTKD AANVAKGAMQGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTA LAKEAJQGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTAKTVLTG TKDTVCSGVTSANNVAKGAVQTGVDTA ANVAKGANVAKGAVQGLDTTKSVLTGTKD ANVAKGANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSANNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGAVQTGM OTTKTVLTGTKDTVTTGLVGAVNVAKGAAQTQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAKGAA	1		ł	1 1			
NTI.KRHGVHRGDRVAIYMPVSPLAVAAMLA CARIGAVHTVIFAGFSAESI.AGRINDAKCKVV ITFINQGLRGGRVVELIKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLTYSGSTIGMFKGIVHT QAGYLLYAALTIHKL VFDHQPGDIFGCVADIG WITCHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQPYGAPTAVRLLLKYGD AWVKKYDRSSLRTILGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPFKPKG KTLGSFFGSLFGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTKAVLTG TKDTVSTGLTGAVNVAKGTVQTGVETS KAVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTQTGVDT AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD VSGVTGAMNVAKGTQQTGVDTT TKDTVCSGVTSAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDT KTVLTGTKDTVTTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTTVTTGLVGAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAKGAA							DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS
CARIGAVHTVIFAGFSAESLAGRINDAKCKVV ITFNQELRGRVVELKKIVDEAVKHCPTVQH VLVAJRTDIKVHIMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVOFQPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTFRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAQQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVAMSAFTOLTGAVNVAKGTQAGVDTTKTV LTGTKDTVTTGTKDTVCSGVTGAMNVAKGT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT VGAMNLARGTIQTGVDTTKTVLTGTKDT ANVAKGAMQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGAVQTG GLKTTQNATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAKGA	1						PESVALIWERDEPGTEVRITYRELLETTCRLA
ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPSSARNLVANAHSSARARPA ADPTGAPAALEAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTKSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTKSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTKSVLTGT TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVAKGTQTG GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTQT GVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKNTVCSGVTGAMNVA KGTIQTGVDTTKNTVCSGVTGAMNVA KGTIQTGTGVDTTKNTVCSGVTGAMNVA KGTIQTGVDTTKNTVCSGVTGAMNVA KGTIQTGVDTTKNTVCTGTKDTVC SGGVTGAANVAKGAVQGGLDTTKSV VIGKDTMSTGLTGAVNVAKGAVQGGLDTTKSV VIGKDTMSTGLTGAVNVAKGAVQGGLDTTKSV VIGKDTMSTGLTGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTTVTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTTVTGLVGAVNVAKGAVQTG GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKNTTVTGLMGAVNVAK	1			i i			NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKL VFDHQPGDIPGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADTGAPAAEAAQPQAQVAAHPEGTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVAKGTQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNLATGTKDTVCSG VTGAMNLARGTIQTGVDTTKSVLTGTKDT CSGVTGAANVAKGAVQGGLDTTKSV UGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK							
VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKL VFDHQPGDIFGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKFKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSG VTGAMNLARGTIQTGVDTTKTVLTGTKDTVCS GVTGAMNLARGTIQTGVDTTKTVLTGTKDTVCS GVTGAMNLARGTIQTGVDTTKTVLTGTKDTVCS GVTGAMNLARGTIQTGVDTTKSVLTGTKD AVSTGLTGAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTTTGLYGAVNVAKGAVQTGM DTTKTVLTGTKDTTTGLYGAVNVAKGAVQTG GLKTTONIATGTKDTTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAKGAA							•
QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSTVVYGPLCNGATSVLFESTPYYPNA GRYWETVERLKINGPYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTSKALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVAKGTQTGV TSKTVLTGTKDTVCSGVTGAMNVAKGT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAJQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNLATGTKDTVCS VTGAMNLARGTIQTGVDTTKTVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSV VGTKDTMSTGLTGAANVAKGAVQGGLDTTKSV VGTKDTMSTGLTGAANVAKGAVQGGLDTTKSV VGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTTGLVGAVNVAKGAVQT GLKTTQNLATGTKNTTGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKNTTTGLVGAVNAKGAA							
WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINGFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGYMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAJQGGLDTTKSVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKTVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQTGUDTA KTVLTGTKDTTTGLVGAVNVAKGTVQTGVDT KTVLTGTKDTTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTTGLVGAVNVAKGTVQTGM CTKDTVCSGVTSAVNVAKGAVQTG GLKTTQNIATGTKNTTFGSGVTSAVNVAKGAVQT GCKTTQNIATGTKNTFGSGVTSAVNVAKGAVQT GCKTTQNIATGTKNTTFGSGVTSAVNVAKGAVQT	1						
GRYWETVERLKINGFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNLATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSV TGTKDTVCSGVTSAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTTTGLVGAVNVAKGAVQTG GLKTTQNLATGTKNTFGSGVTSAVNVAKGAVQT GLKTTQNLATGTKNTFGSGVTSAVNVAKGAVQT GLKTTQNLATGTKNTFGSGVTSAVNVAKGAVQT GLKTTQNLATGTKNTFGSGVTSAVNVAKGAVQT							
AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT 1274 2624 A 10017 I 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSWMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTYTTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTYTTGLVGAVNVAKGAVQTGM GILKTTQNIATGTKNTTGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAKGAA	1						
1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGUTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTK TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGAVQTG GLKTTQNIATGTKNTFGSGVTSAVNVAKGAVQ GGLDTTKSVVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK]		i				
1274 2624 A 10017 1 3750 FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYGLVGAVNVAKGAVQTG GLKTTQNIATGTKNTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTTGLMGAVNVAK							
KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDL VCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGAVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGAVQTGM GLKTTQNIATGTKNTFGSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAVQT	1274	2624	A	10017	1	3750	
ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTTYTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQTG GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK			- "		_		
KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYTGLVGAVNVAKGAVQT GLKTTQNIATGTKNTTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK							
VSSGVTGAMDMAKĞAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGAVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQTGUDTTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTVTTGLVGAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK						1	
TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSNVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQTGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYTTGLVGAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK	1					}	GVASVVDVAKGVVQGGLDTTRSALTGTKEV
LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYTGLVGAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK							
KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK					Ì		
DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK	1			}	1	ļ	
GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMYMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK					l		
IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK			- 1		ſ		
KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK			l				
LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK							
AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK			1		ĺ		
VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLIGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK							A AND A COMMOTCH ATTOM A TOTAL TOTAL
SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK							VTCAMOI ARCTIOTCVDTTVDA TCTVDTV
AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK] [·	- 1	[. [
TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK					ļ		AVSTGLTGAVNVAKGTVOTGVDTTKTVITG
VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK]				•		
KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK			ĺ	l	[ĺ	
DTTKTVLTGTKDTIYSGVTSAVNVAKGÅVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK	1 1	1	ł	į		}	
GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK					İ		
		Į.	ļ		i		GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA
		1	ł	1	1	1	
	L						GTVQTSVDTTKTVLTGTKDTVCSGVTGAAN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
		1		1		
cotide	seq-	ŀ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	j	J	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	[1	[peptide	1	/=possible nucleotide deletion, \=possible
1	i .		1	sequence		nucleotide insertion
	 	 	 	1 5545555		VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA
i '	İ	ļ				VKLAKGTVOTGMDTTKTVLTGTKDAVCSGV
í	ĺ	[1	!	f	
1			l .		•	TGAANVAKGAVQMGVDTAKTVLTGTKDTV
		ł				CSGVTGAANVAKGAVQTGLKTTQNIATGTK
•			1			NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT
1 1	ĺ	1	ı	ĺ	ſ	GTKDAVSTGLTGAVNLAKGTVQTGVDTSKT
1			1			VLTGTKDTVCSGVTGAVNVAKGTVQTGVDT
	i]		1	AKTVLSGAKDAVTTGVTGAVNVAKGTVOTG
[1	l	1		[VDASKAVLMGTKDTVFSGVTGAMSMAKGA
		1	1	i	ł	VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA
1		ļ.	ļ	1		1 3
i			1			TGATHTGLSTFQNWLPSTPATSWGGLTSSRT
1		i		•	Į.	TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP
1		l	ŀ	}	1	AWEAAATTKGLATDVATFTQGAAPGREDTG
]			ļ	J ']	LLATTHGPEEAPRLAMLQNELEGLGDIFHPM
				ł		NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD
]	Į			LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL
j l		1	l	ļ		ODCFRL
1275	2625	Á	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM
12/3	2023	(^	10025	124	413	
		i			•	YKRKYSAANTKVEKKKKEKVLAPVTKPVGG
1		,			ļ	DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG
						KKPFS
1276	2626	A	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS
!				·		ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED
1						EAVDVTPVMTCVFVVMCCSMLVLLYYFYDL
		[[!		LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK
i l			(CRIPNNSLPYFHKRPQARMLLLALFCVAVSV
			ļ			VWGVFRNEDO
1277	2627	Α	10035	51	869	
1 12//	2027	Α	10055	51	809	YSRFTVPLPATMASSEVARHLLFQSHMATKT
1			!			TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL
			i .			PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT
!						LFNTNFEDYESSHFCPNVKLKAQTYELQESN
1			1	İ		VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN
}			ł			PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE
1						SSGQSAQQPYLPINSPPHRLADVADVHLFSSV
		ľ				LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS
1278	2628	 	10036	3	457	KGDQEQGSWKHGADPLRGGEM
14/8	4046	Α	10030	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST
ļ						VMGAVGESLSVQCRYEEKYKTFNKYWCRQP
						CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL
		ļ ļ]		TFTVTLENLTADDAGKYRCGIATILQEDGLSG
						FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR
		· -	1			SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA
1200	2620		10040		244	QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS
						FLRGLFGGNTRIEEACEMYTRAANMFKMAK
[1		ĺ	[NWSAAGNAFCQAAKLHMQLQSKHDSATSFV
		ì		ì	ľ	DAGNAYKKADPQGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG
						PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY
1202	2032	^	10004	-	1040	
1		i				NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA
		i			j	ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF
	ł	I		ľ		YEFQLTAVSEGGVLSESSSTANITVVASDSPY
		j		ļ	·	GRFAFSHEQLRVSEAQRVNITTIRSSGDFGHVR
						LWYKTMSGTAEAGLDFVPAAGELLFEAGEM
		ļ		j		RKSLHVEILDDDYPEGPEEFSLTITKVELOGR
		l		ļ	ŀ	
		J	J			GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN
						AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII LTIYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL ALEGPLLITFFVRRVKGTFGEIM
1283	2633	A	10088	316	516	MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM HLXRS
1284	2634	A .	10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	A	10100	1	574	RPRGRGAWAGPGGDYSGVRRQORRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	A	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1	478	MEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	A	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	A	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLOPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1	749	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \mathbb{\mathbb{-}possible} nucleotide insertion LGNVLTSTPNAKTVNGKAESSDSGAESEEEE AC
1293	2643	A	10124	2	989	PLMSLVRVVEFVAASSAQKTPSRLENYYMVC KADEKFNQLVHFLRNHKQEKHLVFFRYSSGL CGRGIRDSARMCSTCACVEYYGKALEVLVK GVKIMCIHGKMKYKRNKIFMEFRKLQSGILV CTDVMARGIDIPEVNWVLQYDPPSNASAFVH RCGRTARIGHGGSALVFLLPMEESYINFLAIN QKCPLQEMKPQRNTADLLPKLKSMALADRA VFEKGMKAFVSYVQAYAKHECNLIFRLKDL DFASLARGFALLRMPKMPELRGKQFPDFVPV DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN EGRRKFIKNKAWSKQKAKKK
1294	2644	,	10129	91	1042	VTMYKDCIESTGDYFLLCDAEGPWGIILESLA ILGIVVTILLLLAFLFLMRKIQDCSQWNVLPTQ LLFLLSVLGLFGLAFAFIIELNQQTAPVRYFLF GVLFALCFSCLLAHASNLVKLVRGCVSFSWT TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN MTPCQLNVDFVVLLVYVLFLMALTFFVSKAT FCGPCENWKQHGRLIFITVLFSIIIWVVWISML LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLLNSDFLSHSGA
1295	2645	A	10133	376	518	RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG GRGAPRQEGPGSSWRQV
1296	2646	A	10135	3	551	EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGEGSPVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKIIAIFGLEGLGLRDLWKPGIELLQE
1297	2647	A	10138	48	407	MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC
1298	2648	A	10156	94	453	GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT
1299	2649	A	10161	1	393	PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEBAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV
1300	2650	A	10162	98	391	AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV
1301*	2651	A	10165	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV IMEGVQEETDTRDVKRQVERSEICTEEPQKQ

SEQ ID No. of nucl- order peptide conide graph of the period to the period of peptide conide graph of the period period of the period of peptide conide graph of the period of period of the period of p							
nucleotide seq- uence unice un	SEQ ID			SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide seq- wence unice to distantion of the propriet sequence when the propriet sequence with the	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid
Doublow Decision	nucl-	peptide	l	in	, ,	location	
uence 99496 corresponding on the first of the piptide residue of peptide sequence continue to the peptide sequence		1	1			I	
unice and to first amino acid residue of peptide peptide peptide sequence s			1	1			
mino edid reduce of pepide sequence pepide sequence pepide sequence pepide sequence pepide sequence pepide sequence pepide sequence pepide sequence pepide sequence pepide sequence pepide sequence seque	1 .		l	l .			
residue of peptide sequence Poptide in uncloid delicini, "possible nucloid delicini	dence	Ì	[914			
peptide sequence Sequence	į	1	ŀ			1	
Inteledide insertion KSTILNSKHILKKODSETPHLKSILLKKEVKSS KEKPERKITSEDKI SVRHKYKODCHHKTO DETELHSEKGIAL VERNIQKOSOOTKISDDK TERKSKHRINERKI SVIJGKOGKPVSEVIIKTDE NYKERNIKKERILSKIJGKOGKPVSEVIIKTDE NYKERNIKKERILSKIJGKOGKPVSEVIIKTDE NYKERNIKKERILSKIJGKOGKPVSEVIIKTDE NYKERNIKKERILSKIJGKOGKPVSEVIIKTDE NYKERNIKKERILSKIJGKOGKPVSEVIIKTDE NYKENIKKERILSKIJGKOGKVSEVIIKTDE NORSHLKPEVVHREKRITSKILLEGKLIVLKS KSKTQGKOVKVETELQEGATKOATTEKPD KEKNTEENDSKIGKRSKIKHSKISKULKSKIJKKO DEKSTEENDSKIGKRSKIKHSKISKIJKSKIDKKD KSTSTRILERKLSOGKIKSRIKHSKISKIJKKING KSKTQGKOVKVETELQEGATKOATTEKPD KKSDDKLOGKEVDSSHBKARGHIAKEKYKSD NOSTITELERKLSOGKISSISHASKULHAEKYKSD KOSILORTRICEROSSAGHERIKANISISHEKKI, SRILCERRGSLSQEMAKOEKKLAAHTLSTT SGOSSLORPKKSGDMTILDEOEPMEIDSEPOVE NOFEVSKTOPINNINSKIQDDISSEMKOKITS ATVQKDELRTCTADSKATAPAYKPGGGTGV NOFEVSKTOPINNINSKIQDDISSEMKOKITS ATVQKDELRTCTADSKATAPAYKPGGGTGV NOFEVSKTOPINNINSKIQDDISSEMKOKITS ATVQKDELRTCTADSKATAPAYKPGGGTGV NOSUSKITAKHORISOKOMINSKIDDISSEMKOKITS ATVQKDELRTCTADSKATAPAYKPGGGTGV NOSUSKITAKHORISOKOMINSKIDDISSEMANINSKIDDISSEMANINSKIDDISSEN LSSVTVVPLRESVDPVPLFJOKKIVLDSSTA SISSADHAJAPNOSLIVASESEVLKTSSSKEGG EGFTVDTPAKASITSKRIPPEAHQATILDGKO GKVIMPLGSKLTGVVTESEVLKTSSSKEGG EGFTVDTPAKASITSKRIPPEAHQATILDGKO GKVIMPLGSKLTGVVTENENDKYGGLVDMA KKENDLINAEPILKQTIKATVENGKKDGIAVD HVGLINTERYAETVKLEKRKSPGKVKIDSID VERRIENSEVDTSAGSGGAPSVLHQRNOQTE DVATORPRAEKTSVATTEGKDKDVTLSVVK AGPATITISSETROSEVALPCTISIEADBGLIGT HSRNNPLHVGAEASECVTVFAAEEGGAVTE GFASSTFILTTISEGGSGCAVAPSEDRAADL LAVIAVKIERAVNISVTVEKENDAVTSAGSEE KCDGSLSRIBSEIVEGTTTTISEVESGAVVTET GFASSTFILTTISEGGEGGAVTTEGGAVTTE GFASSTFILTTISEGGEGGAVTTEGGAVTTE GFASSTFILTTISEGGEGGAVTTEGGAVTTE GFASSTFILTTISEGGEGGAVTTEGGAVTTE GFASSTFILTTISEGGEGGAVTTEGGAVTTE GFASSTFILTTISEGGEGGAVATSEGDAVTSAG TITAGGISSEEVOGGSOMMMMMGKKKETEG TVTCTGAGEGSDIPVCSVTGAGGFECTISTTGCKGSKDT GGGSAVTSTGGTEGGGAVTTEGGAVTTE GFASSTFILTTINGTGGGSGGAASTGGGAVTOT GGGSAVTSTGTTEGGEGGAVTTEGGAVTTE GFASSTFILTTINGTGGSSGGAASTGGARGGGGT NORTOGGATGGGGATATTAGGGTST ANGGGGATGATAGATTRACAGGST ANGGGGATAGATAGATAGGGST ANGGGGATAGATAGATAGGGST ANGGGGATAGATAGATAG	ļ	ł	1	}		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
KSTILNERHILKEDDSEPHILISLIKEFVISS KEKPERKITSEKOL SVHRIVKODOMIKITO DETELHISSEKOLK VERNIQUOSQOTKL SISDDIK TERK SKIRNERKIL SVIGKORGVISVETYIKTDE NWAKENNIKKERRIL SAEKTRA-BHIKSRRSIDSIS (QKOSLOGKOJEKOLMOST) NMDSNILKPERVVHKEKRITIKSIL EERIL VILKS KIKOTGOKQVIKVETEL DEGATIKOATTIKPD KEKNTEENDISKORK SKVEDIKPEETOVEPV LETASSSAHTOTGONSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEYKYSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEYKYSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKAD DIKSDADOKOMENDA SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIM	1	1	1		peptide		/=possible nucleotide deletion, \=possible
KSTILNERHILKEDDSEPHILISLIKEFVISS KEKPERKITSEKOL SVHRIVKODOMIKITO DETELHISSEKOLK VERNIQUOSQOTKL SISDDIK TERK SKIRNERKIL SVIGKORGVISVETYIKTDE NWAKENNIKKERRIL SAEKTRA-BHIKSRRSIDSIS (QKOSLOGKOJEKOLMOST) NMDSNILKPERVVHKEKRITIKSIL EERIL VILKS KIKOTGOKQVIKVETEL DEGATIKOATTIKPD KEKNTEENDISKORK SKVEDIKPEETOVEPV LETASSSAHTOTGONSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEKYKSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEYKYSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKKD DIKSDADOKOKEVDSSIPHALIPLA-KEYKYSD KOSTSTRILERKI SDGHKSRSI-KHISKODIKKAD DIKSDADOKOMENDA SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIMIKOKTS ANTOKODIKANI SIPHALIPLA-KISTODIM	i	ŀ		1			nucleotide insertion
KEKPERKTTSEDKLSVKHKYKGDCMHKTO DETELHSSEKGLAVENDKOSQOTKLSSDDK TERKSKHRNERILSVLGKDGKPVSEVIIKTDE NYKERNIKKERRILSVLGKDGKPVSEVIIKTDE NYKERNIKKERRILSSVLGKDGKPVSEVIIKTDE NYKERNIKKERRILSSKLGKEKTKSLLEKSKLVLKS NENNIKCEVEVHAKERRTISLLEGLATKOATTEKED KEKNTERDISSKGRKSKVEDKPETETGVEPV LETASSSAHTOKDSSBFAKL PLAKEKYKSD KEKNTERDISSKGRKSKVEDKPETETGVEPV LETASSSAHTOKDSSBFAKL PLAKEKYKSD KOSTSTELERILSDGKESSLKHSKSDKIKKD ENKSDDWDCKEVDSSHBKARGNSSLMEKAL SRRLCENRGSLSQEMAKGEKLAANTLSTP SOSSLQRYKKSGDMTLFQEPMEDISETGVEV NYFEVSKTQDNRNNSHQDIDSENMKQKTS ATVOKGDELRICTADSSETGVEV NYFEVSKTQDNRNNSHQDIDSENMKQKTS ATVOKGDELRICTADSSETGVEV NYFEVSKTQDNRNNSHQDIDSENMKQKTS ATVOKGDELRICTADSSETGVEV NYFEVSKTQDNRNNSHQDIDSENMKQKTS STSPADHSALPNOSLTVRESEVLKTSDSSEGGG GEFTVTVTLESSTVPLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG GEFTVTDTTAKASTISKRHIPEAHQATLLGGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG GEFTVTDTTAKASTISKRHIPEAHQATLLGGSTA KKENDIDAMEPILKGTIKATVPNOKKRGIAVD HVVGLNTEKYABTVKATSTEGKDKDVTLSPVK AGPATITTSETRGSEVAHLPSTISEAGGLIGT HSRNNPLHYGARSGSRSPNLHRONGOTE DVATOFRRAEKTSVATSTEGKDKDVTLSPVK AGPATITTSETRGSEVAHLPSTSEGGLIGT HSRNNPLHYGARSGECAVATSEGGLIGT HSRNNPLHYGARSGECAVATSEGGLIGT HSRNNPLHYGARSGECAVATSEGGLIGT HSRNNPLHYGARSGECAVATSEGGLIGT HSRNNPLHYGARSGECAVATSEGGLIGT HSRNNPLHYGARSGECAVATSGERAADL LAVHAVKIEANVNSVYTEEKDDAVTSAGSEE KCOGLSRDSIFICETTITISEVESDAAVTSAG TERKAGSISSEEVOGSOGNMMRMOFKKETEG GESAVTSTOTASGEGEGGAVTTEG GESAVTSTOTASGEGEGGAVTTEG GESAVTSTOTASGEGEGDVATTORGEGIGH ASTCTGLÆGERSDBVYTTISTEVESDAAVTSAG TERKAGSISSEEVOGSOGNMMRMOFKKETEG TYCTGAGERSDBVTSTAKSGETAVIDLAGGG GYMTSAASDQSDQLEKVEDTTISTGVEGGG VTVLTSGEGEGDVATSTAGREGIGH ASTCTGLÆGERSDBVTANDVSNSTEKERGSKDT DICSSAKGIVESVTSAVSGKDEVTPVPGGG GPMTSAASDQSDQLEKVEDTTISTISTEVESDAA LAVHAVERARNYSVYTEKENDAAVSGEDSUTDSTE REGGEAVMIGAVTAGGEGEPLYSSATTISTEGSKDT TYCTGAGERGBDEGGAVTSEGGEDGVATTSGEGGH ASTCTGLÆGERSDBVTTARSFSEKEDDDISTISTE NEGGEAVMIGAVALORDAAPFSASGEGGSGVATGNEIGH ASTCTGLÆGERSDBVTANTORNOSTSTERVELC GEGAVTSTOTAGEGEGGEAVTSTSAGGGGTANTSTATICAGGS TYDAVASGERPECAVATSGEGEPLISSATICAGGSS GVYVESENERGTVANTERVENDELDISTNE GEGEAVMIGAVALORDERDISTSSVECC EGGVSSAVPGEGG		 	-	t			
DETELHSSEKGILKVEENIGKQSQOTKLSSDIS TERKSKHENERKI.SVIGKDGRYSEYSEVIKTIDE NVEKENNIKEERI.SAEKIKAEHKSRESSISS (IOKDSILGSKQHGTICQRSESYSEDKCDMDST NMDSNILKPEEVVHKEERERTKSILLEEKILVKS SKORGKQVKVVETELQGGATKATPKPD KEKNTEENDEKGRISSHEAKILEEKILVKSP KEKNTEENDEKGRISSHEAKILEEKILVKSP KEKNTEENDEKGRISSHEAKILEEKILLEEKILVKSP KEKNTEENDEKGRISSHEAKILEEKILLEEKILSDHKSRISKKUK DENESSAHISTOKUSSHEAKARGISSIMEKKIL SKRILCENRRGIS.GOBMAKGEKILAANTI.ST SRILCENRRGIS.GOBMAKGEKILAANTI.ST SRILCENRRGIS.GOBMAKGEKILAANTI.ST SOSSLORPIKKSODMTI.PROPEMBESPETOVE NVFEVSKTODNRNNINSHODDSENMEKORT. ATVQKDELRICTADSKATAPAYPGRGITOV NSNSEKHALDHRST.TIKKAHIQSAVSKMINGE KEPHRGTTEVNIDSSTVIRMLESSTVA. ATVQKDELRICTADSKATAPAYPGRGITOV NSNSEKHALDHRST.TIKKAHIQSAVSKMINGE KEPHRGTTEVNIDSSTVIRMLESSTVA. STSPADHSALPYQGI.TVRESEVIKTSDSKEG GEFTVOTDTAKASTSIKHIPPAHQATILDGKQ GKVMPPLOSKLTOVIVENERITKEGGI.VOMA KKENDINABENILKOTTIA.TVRESEVIKTSDSKEG GEFTVOTDTAKASTSIKKHIPPAHQATILLDGKQ GKVMPPLOSKLTOVIVENERITKEGGI.VOMA KKENDINABENILKOTTIA.TVRENGKLOGIAV HVGILNTEKVAETVKLKHKRSPOKVKIDISID VERRNENSEVDTSAGSGSAFSVLHQRNIOGTE DVATOPRRAKETSVATSTECKOKOTVLSTOW AGPATITISSETRQSEVALKOTSIELDEGLIOTT HSRNIPHLYGGALSSECTVFAAAEGGAVVTE GRAESITH.TSTKEIRGESGCAVAESEDRAADIL LAVIHAVKIRANNISVVTEEKDDAVTSAGSEE KCDGSLSRSDERVEGTITITISEVBSDGAVTSAG TERAAGSISSEEVVBGSOONMMEMRGYKETBG TVTCTGAGERSDNIPVICSVTGAGRREERMYT GAGVVLGINDAPPGTSAGSGGOSSEGTANTOTT GGSAVTSTGTEDGEGPASCTGSEDSSEGTAS SESSEMGSAMDSTVAKSTOTTSTGCHUGGI GGSAVTSTGTEGGEGEPATYTSTGCHUGGI GSAVTSTGTEGGEGEPATYTSTGCHUGGI GGSAVTSTGTEGGEGEPATYTSTGCHUGGI ODGGIVTSTGAKEEDEGEBOVYTSTGRGREEGI TVTCTGAGESSENGCHLESKAGONGOTTOTTSTG. GGSAVTSTGTGGGEGESGOSGGANTOTTSTGCHUGGI ASTCTGGGESSEGVALASTGEGGSOSGGANTOTTSTGCHUGGI GGSAVTSTGTGGGEGSTSSEGGAL SESSEMGEAMDSTAAGGTNYNYNAGGFCCGGSOSGGANTATICKAES LQPVAAAVERATOFYLISTADFEGRIPTSYVE NECCDGLMATTASGOITNONSLAGGKNOCK VIJISTSTNDYTTYGYGGCE GGMTSAASDQSDSQLEKVEDTTTSTGL.VGGS GVVYSENERRAGTVMEEKDGGGISTSSVEGOL GERSAVTSGGGEGGAMDSTANTEGEBOTTSTTATCE EMMEGTEVTTEFFANPFANGSGDSGGTSSSEGGA GVVYSENERRAGTVMEEKDGGGISTSSVEGOL GGLSATEVSKHVPAPSALJAERNCRCFGGFV GGKEFGGFANSTATTREGGISTSSTOOL GGLSATE	1		1				
TERKSKHRINEKLISVLGKDGRYVSEVIILDE WYRKENNKKERERISABEKTASHENSENSE IQKDSLOSKQHGITLQRRSESYSEDKCOMDIS IQKDSLOSKQHGITLQRRSESYSEDKCOMDIS IMDISNIK/PEEVVIKKERRITKSLLEEKLVLKS KSKTQGKQVKVVETELQEGATKQATTREPD KEKNTEENDESKQRSKVENDER/PEETGVEPV LETASSSAHSTOKDSSIRAKLPI-AKEKYKS KSKTQGKQVKVVETELQEGATKQATTREPD KEKNTEENDESKQRSKVENDER/PEETGVEPV LETASSSAHSTOKDSSIRAKLPI-AKEKYKS KSKTQGKQVKVVETELQEGATKQATKEKKD ENKSDDKDGKEVDSSIEKARCHSSLMEKKL SRRLCENRRGSLSQEMAKGEEKLAANTLST SGSSLQRPKKSGDMTLFEQEPMEDISEROVE NYFEVISITQDNRNNNSHQDDSENMEQKIT ATVQKDELRICTADBKATAPAYPERGITGV NSNSEKHADDIRSTLTKKHHIQSAVSKMNFG EKPHRIGTTEVNIDESTVFIRMLISAFSENDER QKNLKNTAAEEHVAQODATLEHSTNLDSSTS STSPADHSALPNQSLTVRESEVLKTSDSKEEG EFTYTOTPAKASTISKRHPEARJATLLDGKQ GKVIMPLGSKLTGVVENENTIKEGGLVDMA KKENDLNAEPILKQTTLAVTLOKKTOGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRHENSEVDTSAGSGASPSVLHQRNQQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETTQSGVAJECTISTERE EFTISTISTEGSGEAVAREGGGAVTE GAESEFFILTSTIKEGSGSGASVSVLHQRNQQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETTQSGVAJECTISTESTEGSGAAVSAGA TERRAGSISSEEVDGSQGNMMEMGPKKETGG TOTTCTGAGERSIDNYICISTTGAGTGETMAT GAGVVLGDDAPPGTSASQEGDGSVNDGTE GESAVTSTOTTCGAGESGSTVTASAGEGGAVTE GESAVTSTOTTCGAGESGSTVTASAGEGGAVTE GESAVTSTOTTCGAGESGSTVTASAGEGGAVTE GESAVTSTOTTCGAGESGTVLICESAGGAGGAGGAT TERRAGSISSEEVDGSQGNMMEMGPKKETGG TOTTCTGAGERSIDNYICISTTGAGTEREMYT GAGSVLGDDAPPGTSASQEGDGSVNDGTE GESAVTSTOTTCGAGESGTVLICESAGGAGGAGGAGGAT SISSEENGESAMDSTVAKEGTNYPLVAAGFGEGH ASTCTLGAGESGGVLICESAGGOGGTVVDTET GESAVTSTOTTCGAGEGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA			l	ļ			
NVRKENNKERRLSAEKT/AEHKSRRSSDS IOKDSILGSKOHGTICRRESEYSEBOCDMDST NMDSNLKPEEVVHKERRRTKSLLERKLYLK KSKTOGKOVKVVPTELGGGGATKACATTEKPD KEKNTEENDESKORKSKVDDKPTELTGVEPV LETASSSAHSTOKDSSITRAKLPLAKETYKSD KEKNTEENDESKORKSRVALPLAKETYKSD KEKNTEENDESKORKSRVALPLAKETYKSD KEKNTEENDESKORKSRVALPLAKETYKSD KEKNTEENDESKORKSTAKLPLAKETYKSD KOSTSTRLERKLSDOHKSRSLKHSKNIKKKD ENKSDDKDGKEVDSSITRAKLPLAKETYKSD KOSSLORPKKSODMILPEGEMEDISEDYCY NYFEVSKTODNINNISHODDISENMEQKITS ATVQKDELRTCTADSKATAPAYKPGRGTOV NSNESKHADHRISTIKKMHIQSAYSKMNFGE KEPHRGTITEVHIDSETVHRMLSAPSENDSP UKNINTAAEEHVAQODATLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLSOSSGG KEPHRGTITEVHIDSETVHRMLSAPSENDSPS LSSVTVVPLRESYDPDVPLFDKRTVLSOSSGG GGFTVDTPAKASITSKRHIPPAHQATLLDGKQ KWIMPLGSKLTGVVVENENTIKEGUVDMA KKENDLNAEPNLKOTIKATVENOKKDGLAVD HVGINTEKVAETVIKLKHRSPGKVKDIDSID VERRHENSEVDTSAGSGARSVLHORNOGTYE DVATOPRABKTSVATSTECKKDVTLSPVX AGPATITISSETRGSEVALPCTSIEADEGLIGT HSRNNPHHVGABASECTVFAAAEEGGAVVTL GGAESETPLISTIKEGESGECAVAESDGAVADL LAVHAVKIRANVNSVYTEIEKDDAVTSAGA TEIRAGSISSEEVDGSOGMMREMGFKERTEG GGASVTLGINDAPPGTSASGGBOCSVNDGTE GGASVTSTGITGGGEGGPASCTGSSDSGGALS SESSENGESAMDSTVAKEGTNVILVAAGGCE GGAVTLGINDAPPGTSASGCOSSNSGGTAIS SESSENGESAMDSTVAKEGTNVILVAAGGCE GPMTSAASGOSDSCLEKVEDTTSTGLVGGGE GPMT	1		[1		İ	
IQKDSLGSKQHGITLQRRSESYSEDKCDMDST NMDSNLKPEEVVIKERRITKSLEEKLVLKS KSKTQGKQVKVVETELQEGATKQATTIKPD KEKNTENDSKGKKSKVEDKPFETGVEPV LETASSSAHSTQKDSSHRAKILAKEKYKSD KDSTSTRLERKLSDGHKSRSLKHSSKDIKKKU BIKSDKLGKEVVDSSHEKARGNSSLMEKKL SRRLCENRRGSLSQEMAKGEKLAANTLST SCSSLQRPKKSGDMTLFPGQFMEDISPROYE NVFEVSKTQDNRNNNSHQDDSENMKQKIT ATVQKDELRTCTADSKATAPAYPGGGTOV NSNSEKHADHRSTLTKKMHQSAVSKMNGG KPHKRGTTEVVIDSSTVPRMLSASSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSTS LSSVTVPLRESYPDVPLPLDKRTLLGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG GGFTVDTPAKASTSKHPPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVV HVGGLNTEKYAEFVLKLKKRSPKGKVGDISD VERRNENSEVDTSAGSGARSVLHQRNQGTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGIT HSRNNPLHVGAEASECTVFAAAEGGAVVTE GFASSETFLTSTKEGESGECAVAESEGRAADJ LAVHAVKIEANVNSVVTEKDDAVTSAGSE KCDGSLSRDSEVEGTITFISEVSBOAVYTSAG TEIRAGSISSEEVDGSQGNMMRMGFKKETBG TVTCTGGAGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPGTSASGGGARSVLHOPLOGT GESATSTGTEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNPLVAAGCD DEGIVTSTGAKEGESGGVLICESAGGDSGTVTPVT GGAVALGDNDAPGTSASGGGGGTSWDTT GGESATSTGTEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNPLVAAGCD DEGIVTSTGAKEGDEGDEVATSTGGNEIGH ASTCTGLGESSGVLICESAGGDSGTVTPVT GGCEG GFATSTGTEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNPLVAAGCD DEGIVTSTGAKEGDEGEDVATSTGGNEIGH ASTCTGLGESSGVLICESAGGDSQGTVTET GESAAGTUSGSAGGDVAKEGTNPLVAAGCD DEGIVTSTGAKEGDEGDEVATSTGGNEIGH ASTCTGLGESSGVLICESAGGDSQGTVATE GESAATSTGTEDGEGBOVTSTGGNEIGH ASTCTGLGESSGVLICESAGGDSGGTASS SESEENGESAMDSTVAKEGTNPLVAAGCDC DEGIVTSTGAKEGDEGDEVATSTGGNEIGH ASTCTGLGESSGVLICESAGGDSQGTTVETPUGGCG GPMTSAASQCSDSQCLEXCEDTITSTGLVGGGS YDVLVSGEVESVTSAVSGCDEVTPVTGGCE GFANSASCDSSDGALGSTGSTSVEV NECCGGLAMTTAGGGVTGAGGKTGGA VLUSTGSTTADYTPQVSAITDVFGGGGAGGAGGAAGAANANAENNYDSMSGTAGGKTGGG GPMTSAASGCBGGAGGAAGGAGGAGAAANAAGAAGGRAGGAGGAGAAANAAGGRAGGAGGAGAAANAAGGRAGGAGGAGGAGAAANAAGGAGGAGGAGGAGGAGGAGGAGG	1						
NMDSNILKPÉEVYHKÉRERTKISLLEEKLYLK KSKTOGKOVKVYETELGGGATKACATTPKPD KERNTEENDEKORKSKVEDKPFETGYEVEY LETASSAHISTOKDSSIRAKUPLAKEKYKSPU KERSSAHISTOKDSSIRAKUPLAKEKYKSPU KERSSAHISTOKDSSIRAKUPLAKEKYKSPU KDSTSTELERKLSDGHKSRAKHSSKDIKKAC ENKSDBKOGKEVPOSSIEKARGHSSKDIKKAC ENKSDBKOGKEVPOSSIEKARGHSSLMEKKL SRRLCENRRGSLSQEMAKGEEKLAANTLSTP SGSSLORPKKSODMILIPEQEYMEIDSEROW; NYEVSKYKODNINNSHODIDESIMSKOKTS ATVOKDELRICTADSKATAPAYKPOKGTOV NSNSEKHADHRSTITKKMHIQSAVSKMNPGE KEPHRGTITEVYHDSETVHRMLLSAPSENDSKY OKNLKNTAAEEHYAQODATLEHSTNLDSSPS LSSVTVVPLRESYPPDVIPLEDKRTVLESSPS LSSVTVVPLRESYPPDVIPLEDKRTVLESSEGG EGFTVDTPAKASITSKRHIPBAHQATLLDGKOK GKVIMPLGSKLTGVIVENENTIKESDSKEGG EGFTVDTPAKASITSKRHIPBAHQATLLDGKOK GKVIMPLGSKLTGVIVENENTIKEKPOKATOSIGU VERRHENSEVDTSAGSGARSVLHORNOTOTE DVATOPRABKETSVATSTECKOKGVISID VERRHENSEVDTSAGSGARSVLHORNOTOTE DVATOPRABKETSVATSTECKOKGVISID VERRHENSEVDTSAGSGARSVLHORNOTOTE HORDITATIVENENTIKETSTERGEVALPCTSIEADEGLIGIC HERNIPHLYOTAGASGARSVLATSTECKOKATOSIGU APPATITSSETRGSEVALPCTSIEADEGLIGIC HERNIPHLYOTAGASGCAPSVLHORNOTOTE HERNIPHLYOTAGASCOTATSTECKOKATOSIGU HAVHAVKEANVASVTEKENDAVTSAGSE KCDGSLSRDSEEVEGTITTIFISEVESDAVTSAG GFAESETFILTSTEEGESGECAVAESEDRAADL LAVHAVKEANVASVTEKENDAVTSAG TEIRAGSISSEVDGSQOMMRIMMOFKKETEG GFASTTITTIFIGETGGESGCAVAESEDRAADL LAVHAVKEANVASVTEKENDAVTSAG TEIRAGSISSEVDGSQOMMRIMMOFKKETEG GESAVTSTOTEDEGEPASCTGSSESGGAIS SESEENGESAMDSTVAREGTNYPLVAAGPCD DEGITYTTGAEREEDPTSTCRONEGICH ASTCTIGAGERSDYLICESAGGDSSGGGTAVS SESSENGESAMDSTVAREGTNYPLVAAGPCD DEGITYSTGAKEEDEGPAVSTTSGRONEGICH ASTCTIGAGERSDYLICESAGGDSGGTVEVDTG GESAKGTVESSYTSAVSGKDEVTPYPGGGC GFATTSAASDQSDSOLEKVEDTTSTGLVGGS YDVLYSGEVESSYTSAVSGKDEVTPYPGGGC GFATTSAASDQSDSOLEKVEDTTSTGLVGGS YDVLYSGEVESSYTSAVSGKDEVTPYPGGGC GFATTSAASDGSDSGALSKUDTTSTGLVGGS YDVLYSGEVESSYTSAVSGKDEVTPYPGGGC GFATTSTAASDGSDSGALSKUDTTSTGLVGGS YDVLYSGENEERGTOTAMSTS TSECCEAVMIGAVLQUEDGRITTKVEDLSDA AISTSTTADCTPTSASDTREENDLATADREGN GOLSATSVSKHVYMPSLLARNNCRCPGFVY GCGFGPTVANSTEEGGGFGPTAGGSST ASYSAGRGLEGHANSPAHLRGPEGTSGGTAK BUSHSSISHYLAAVNIGAKADAPOGTYA	ŀ	1		1		1	NVRKENNKKERRLSAEKTKAEHKSRRSSDSK
KSKTOGKÖVKVVETELGGATKCATIPRED KEKNTENDSEKQRESKVEDGATKEPETGVEPY LETASSAHSTOKDSSHRAKI,PI.AKEKYKSD KDSTSTRI.ERIK.LSOGHKSKSLIKHSKUDIKKKD ENKSDDKDCKEVDSSHRKAKI,PI.AKEKYKSD KDSTSTRI.ERIK.LSOGHKSKSLIKHSKDIKKKD ENKSDDKDCKEVDSSHRKARGNSSLMEKKL SRRI.CENRROSI.SQENAKGEERLAANTI.STP SGSSLQRFKKSGDMTLIPEGEPHEIDSEPGYE NYFEVSKTODNRNNNSRIQDIDSENAKGKTS ATVQKDELRITCADSKATAPAYKPGRGTGY NSNSEKHADHRSTLTKKMHIQSAVSKMNFGE KEPHRGTTEVHDSETYHRMLI.SAPSENDRY QKNI.KNTAAEEHVAQODATLEHSTNI.DSSPS LSSVTVYLRSSYDPDVPIPTOKRTVLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGGLIVMA KKENDLINAEPILKQTIKATYGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGGLIVMA KKENDLINAEPILKQTIKATYGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTFAKASTISKRIPEGGLIVANA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGAVTTSG GFASETTITTSTKEGESGCAVAESEGRAVITSG GFASETTITTSTKEGESGCAVAESEGRAVITSG KCDGSLSRDSEIVEGTTTTISEVESDGAVYTSG TEIRAGSISSEVEDGSGONMRMGPKKETEG TVTCTGABGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSASGEGGSOSQIGTVVEH VALAGAGAAMNADAVALGTGSBOSSTANDTT GESAVTSTGTEDGEGPASCTGSEDSSEGALS SSESENGESAMDSTVAKEGTITISTULYGGG GFASATSTGTEDGEGPASCTGSEDSSEGALS SSESENGESAMDSTVAKEGTITISTULYGGG GFASAASGOSQIGCAKUNDATTSVE NEGCGAMTSTGTEDGEGPASCTGSEDSSEGANS SSESENGESAMMANDSNAERMONNOTORALGGKNOGK VLIISTSTTNDYTTQVSGGLISTSVEDC GGKATSAASDQSDSQLKKVEDTITSTULYGGG GFASAASGOSQIGCATANDSALAGGRANGATATT ENGGGRAVTGSBORGGGFANSTSVEDC GGKGFGPVLAVSTEEGHLOPSTYRPSAGGDSQITAK DSVACAREEKHOKEGPEIGPFAGROGKESTL HLINAEKNVLLINSLOKEDSKSPTGTAGGSST ASYSAGRGLEGNANSFAHLROP							IQKDSLGSKQHGITLQRRSESYSEDKCDMDST
KSKTOGKÖVKVVETELGGATKCATIPRED KEKNTENDSEKQRESKVEDGATKEPETGVEPY LETASSAHSTOKDSSHRAKI,PI.AKEKYKSD KDSTSTRI.ERIK.LSOGHKSKSLIKHSKUDIKKKD ENKSDDKDCKEVDSSHRKAKI,PI.AKEKYKSD KDSTSTRI.ERIK.LSOGHKSKSLIKHSKDIKKKD ENKSDDKDCKEVDSSHRKARGNSSLMEKKL SRRI.CENRROSI.SQENAKGEERLAANTI.STP SGSSLQRFKKSGDMTLIPEGEPHEIDSEPGYE NYFEVSKTODNRNNNSRIQDIDSENAKGKTS ATVQKDELRITCADSKATAPAYKPGRGTGY NSNSEKHADHRSTLTKKMHIQSAVSKMNFGE KEPHRGTTEVHDSETYHRMLI.SAPSENDRY QKNI.KNTAAEEHVAQODATLEHSTNI.DSSPS LSSVTVYLRSSYDPDVPIPTOKRTVLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGGLIVMA KKENDLINAEPILKQTIKATYGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTPAKASTISKRIPEGGLIVMA KKENDLINAEPILKQTIKATYGATLLDGSPS LSSVTVYLRSSVLAVENSEVLKTSDSKEGG EGFTVDTFAKASTISKRIPEGGLIVANA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGLIVMA KKENDLINAEPILKQTIKATYALGGGAVTTSG GFASETTITTSTKEGESGCAVAESEGRAVITSG GFASETTITTSTKEGESGCAVAESEGRAVITSG KCDGSLSRDSEIVEGTTTTISEVESDGAVYTSG TEIRAGSISSEVEDGSGONMRMGPKKETEG TVTCTGABGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSASGEGGSOSQIGTVVEH VALAGAGAAMNADAVALGTGSBOSSTANDTT GESAVTSTGTEDGEGPASCTGSEDSSEGALS SSESENGESAMDSTVAKEGTITISTULYGGG GFASATSTGTEDGEGPASCTGSEDSSEGALS SSESENGESAMDSTVAKEGTITISTULYGGG GFASAASGOSQIGCAKUNDATTSVE NEGCGAMTSTGTEDGEGPASCTGSEDSSEGANS SSESENGESAMMANDSNAERMONNOTORALGGKNOGK VLIISTSTTNDYTTQVSGGLISTSVEDC GGKATSAASDQSDSQLKKVEDTITSTULYGGG GFASAASGOSQIGCATANDSALAGGRANGATATT ENGGGRAVTGSBORGGGFANSTSVEDC GGKGFGPVLAVSTEEGHLOPSTYRPSAGGDSQITAK DSVACAREEKHOKEGPEIGPFAGROGKESTL HLINAEKNVLLINSLOKEDSKSPTGTAGGSST ASYSAGRGLEGNANSFAHLROP	ŀ	1		į			NMDSNLKPEEVVHKEKRRTKSLLEEKLVLKS
KEKNTEENDSKORKSKVEDKPFETGYEPV LETASSAHTSOKDSSHRAKJAKEKYKSD KDSTSTRIERKLSDGHKSESILKHSSKDIKKKD ENKSDDKDGKEVDSSHEKARGNSLMEKKL SRRLCENRRGSLSQEMAKGERSLAMEKKL SRRLCENRRGSLSQEMAKGERSLAMEKKL SRRLCENRRGSLSQEMAKGERSLAMEKKL SRRLCENRRGSLSQEMAKGERSLAMEKKL SRRLCENRRGSLSQEMAKGERSLAMTLSTP SGSSLQRKKSGDMTLIFEQEPMEIDSETGYE NVFEVSKTQDNRNNNSKIQDDSENNKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTOV NSNSEKHADHRSTITKKMHQSAVKMNPGE KEPHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEHVAQGADTHSILLSAPSENDRV QKNLKNTAAEHVAQGADTHSILLSAPSENDRV QKNLKNTAAEHVAQGADTHSILLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNILKQTIKATVENGKEDGTAV STSPADHSALPNOSIT VRESEVTLEGSTA STSPADHSALPNOSIT VRESEVTLEGSTA STSPADHSALPNOSIT VRESEVTLEGSTA STSPADHSALPNOSIT VRESEVTLEGRAGGAVTUP GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNILKQTIKATVENGKEDGIAV VERRINENSEVTSAGSGLSRAVLHQRRQQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETFRGSSVALPCTSIEADEGLIIGT HSRNIPPLHVGAEASECTVFAAAEEGGAVVTE GFAESFFTLTSTKEGGSGGLGVASEDGRADL LAVHAVKIEANVNSVVTEEKDDAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSAGSGGDSVNDGTE GESAVTSTGTEDGEGPASCTGSEDSSEGTAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEEGBDVVTSTGRGNEIGH ASTCTGLGEESSEVTLGSVAGFBGGGISVNDGTE GESAVTSTGTEDGEGDASCTGSEDSSEGTAIS SESEENGESAVISTGTTEDGEGDASCTGSEDSSEGTAIS SESEENGESAVISTGAKEEDEEGDVVTSTGRGNEIGH ASTCTGLGEESSEVTSAVSGKDEVTPVFVGGG GRATTSAASDQSDSQLEKVEDTTTSTGL VGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTINNSLAGGKNQGK VLIISTSTTINDYTTPQVSAITDVEGGLSDALTRE ENMEGTRYTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTAGGGISTSVEDC GGPVSSAAPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMGGAVLQDDERLTTTEVELDLDA ALISTSTAECMPISAGDRENTSTREDELDGITSSVED GGRAYSTRAERGAVLQDERLITTEVELDLDA ALISTSTAECMPISAGDRAVTPERMAGDTAMISTS TSEGCEAVMGGAVLQDERLITTEVELDLDA ALISTSTAECMPISAGDRAVLTPEDLADARGNESTL ENMEGTRYTTEEFFAPMPSANGGDSQLTAS SESEKDECAMISTIGGEFFELPISSATTIKCAES LQPVSAAVEEREKREPOLTADPEGA GGKEPGVLANSTRILGENDLTADPEGA GGKEPGVLANSTRILGENDLTADPEGA GGKEPGVLANSSLARHERQLTAGGGKSTL HLINAEENVLLINSLIKEBLSKEPTGTAGGSST	i	1		i .		!	
LETASSAHSTOKDSHRALPHAKEKYKSD KDSTSTRIERKLEGHEKSIKOKKKU ENKSDDKDCKEVDSSHEKARONSSLMEKKL SRILCENRROSLOGMAKGERLAANTLSTP SOSSLQRYKKSGDMTLIPEGEPMEIDSEPGYE NVFEVSKTQDNRNNNSHQDIDSENMKQKTS ATVQKDELRICTADSKATAPAYRPGRGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMPGE KEPHRGTTEVHOSETVHRMLASSANDRIG KEPHRGTTEVHOSETVHRMLASSANDRIG KEPHRGTTEVHOSETVHRMLASSANDRIG KEPHRGTTEVHOSETVHRMLASSANDRIG KEPHRGTTEVHOSETVHRMLASSANDRIG KEPHRGTTEVHOSETVHRMLASSASINDRY QKNLKNTAAEEHVAQODATLEHSTNLDSSPS LSSVTVVPLRSSTVPDVPHFAVTLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTDAKASITSKRHIPEAHQATLLDGKQ GKVMPLGSKLTGVIVENENTIKEGGLVDMA KKRNDLNAEPNIKQTIKATVENGKKDGIAVD HVVGLNTEKYASTVKLKHKRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQOTE DVATGPRRAEKTSVATSTEGABUCHLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGGSGCAVAAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSADSETVEGTTTTISEVESDGAVTSAG TERAGSISSEVDGSQGNMRMMGPKKETEG TVTCTGABGRSDNFVICSVTGMREREMYT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTIEDGEGPAASTOSSESGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEGEDVTSTTGRESDSEGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEGEDVTSTTGREIGH ASTCTTGLGEESEGVLICESAGEDSGGTVEH VEAEAGAAMMANENNVDSNASGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GFMTSAASDQSDSQLEKVEDTTISTGLYGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTNINGLAGGKNOGK VLLISTSTTNDYTFQVSATITDVERDLEGALS SEEKDECAMISTISIAEECEASVS GVVVSENERRAGTVMEKKDGGISTSTSVECC EGPVSSAKPVEGEODPSVTPAEEMGDTAMISTS TSEGCEAVMIGGAVLQDEPRITTSGLJAGAGN GALSTSTTAECHPISASTIKEAES LQPVAAAVEERATOPVLISTADEEGPMPSAPP EAESPLASTISKEEKDECALISTISIAEECEASVS GVVVSENERRAGTVMEKKDGGISTSTSVECC EGPVSSAKPUESGGANAVLQDEPRITTSGLJADAGN ALISTSTAECPEIPASASTREEMDOTAMISTS TSEGCEAVMIGGAVLQDEPRITTSGLJADA ALISTSTAECPHISASTIKEBELGHNCRCPGPVR GKEPGPVLANSTEEGHNOPSVERFSAGQGH PSAVCAKEEKHGKECPEIGPFAGROGKSTL HLLDMEKNYLLINSLIKEDENSPHEFRAGGATAK BSSSGRRIGENNANSPAHLRGPEGTSGGTAK DSSVSKRYLAAAVNTGAKKADDMPPVOGTVA	1	1	ľ	i i	'		, , , , , , , , , , , , , , , , , , , ,
KDSTSTRLERKÍ SDGIKSSLKHSSKUKKKU ENKSDDKOKEVDSSHEKARONSSLMEKKI. SRIJOKORKEVDSSHEKARONSSLMEKKI. SRIJOKORKEVDSSHEKARONSSLMEKKI. SRIJOKORNINSKOJOMTLIPEQEPMEIDSEPGVE NVPEVSKTODNRINNSKOJODSENIKOKTS ATVOKDELRICTADSKATAPAYKPORGTOV NSNSEKHADHRSTITKKMIDDSENIKOKTS ATVOKDELRICTADSKATAPAYKPORGTOV NSNSEKHADHRSTITKKMIDSENIKOKY QKNILKNTAAEHVAQODATEVILKISTNLDSSPS LSSVTVVPLRESVPPDVPLFDKRTVLEGSTA STSPADHSALPNOSLTVRESETISTNLDSSPS LSSVTVVPLRESVPPDVPLFDKRTVLEGSTA STSPADHSALPNOSLTVRESETISTNLDSSPS LSSVTVVPLRESVPDVPLFDKRTVLEGSTA STSPADHSALPNOSLTVRESETISTLDGGO GEGTTVDTPAKASITSKRHIPEAHQATLLDGKO GKVMPLGSKLTGVTVPLENETISTGGI.VDM HVVGLNTEKYAETVKLKHKRSPGVKYGLISD VERRINSEVPLTSKAGSOSAPSVLHQRNQOTE DVATGPRRAEKTSVATSTEGKDKDVTISPVK AGPATTTSSETROSEVALPCTSIEADEGLIGT HSRNIPLHVGAEASECTVFAAAEEGGAVTE GFAESETFLTSTKEGGSGGCAVAESEDRAADL LAVHAVKIEANVNSVVTEKEDDAVTSAG SCHENBENDEN GFAESETFLTSTKEGGSGGCAVAESEDRAADL LAVHAVKIEANVNSVVTEKEDDAVTSAG TERAGSISSEVDGSOGNMARMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDSVNDGTE GESAVTSTGTEDGEGPASCTGSEDSSEGTAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGBDVVTSTGRGNEIGH ASTCTGLGEESESVTSAVSKOKDVTPVPVAGGCE GEMTSAASDQSDSOLEKVEDTTISTGL VGGS YDVLVSGEVPECEVAHTSPSEKEDDITTSVE NEECDGLMATTASGDTINNSLAGGKNQGK VLISTSTTINDYTTQVSAITDVEGGLSDALTRE ENMEGTRVTTEEFFAAPMSANSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATITICAES LQPVAAAVEERATOPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTISIAEECEASVS GVVVSSENERAASTOPTLESTADFEGMPSAPP EAESPLASTSKEEKDECALISTISIAEECEASVS GVVVSSENERAASTOPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTISIAEECEASVS GVVVSSENERAANTSLAGGTISTSVEPCO GGKEPGVLANVTSEKARDTAGGNESTL EINMEGTRVTTEEFFAAPMSANGODSQLTAS RSEEKDECAMISTSIGEFEILPISSATITKEDLSDA AIISTSTABCMPISASIDRHENQLTADPEGN GDLSATEVSKHKVPMFSLLARNNCCCGPVR GGKEPGVLANVTSEKARDDARDPVOGTVA HULDHAREKNVLLNSLYKEDENDSPSVHEPSAGQGH PSAVCAEKEEKHGKECPEICPFAGRGQKSETL HLINABEKNVLLNSLYKEDELNDSPSVHEPSAGQCH PSAVCAEKEEKHGKECPEICPFAGRGQKSETL HLINABEKNVLLNSLYKEDELNDSPSVHEPSAGQCH PSAVCAEKEEKHGKECPEICPFAGRGQKSETL HLINABEKNVLLNSLYKEDLANDSCOGTAK DSVSIRVLAAAVNTGAKAADDMPPVOGTVA		1		<u> </u>			3
ENKSDEKDGKEVDSSIEKARGNISSIMEKEL SRRICENRIGGIS.QDMAKGERIA ANTI.STP SGSSLQRPKKSGDMTLIFEQEPMEIDSEPGVE NVFEVSKTQDNRNNINSHQDIDSEMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAYSMNPGE KEPHRGTTEVNIDSETYPIKMLISAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPUFRKILSAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPUFRKTVLEGGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEFNLKQTIKATVENGKKGDGAVD HVVGLNTEKY AETVEKLKHRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRRAKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADGLIGT HSRNNPLHVQLAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMNGPKETEG TVTCTGAEGRSDNPVICSTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEGDEGDEDVTSTGRGREERMVT GAGVVLGDNDAPPGTSASQEGGGSVNDGTE GESAVTSTGTIEDGEGPASTSAGGGGSVNDGTE GESAVTSTGTIEDGEGPASTSAGGGGSVNDGTE GESAVTSTGTIEDGEGPASEAGDSQGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSYTSAVSBEVTYPPGGCE GPMTSAASDQSDSQLEKVEDTTISTSVE NECCDGLMATTASGDITNQNSLAGGNQGK VLISTSTTIADTYPTQVSATITYPSGGLSDALTRE ENMEGTRVTTEFEAPMPSAVSGDDSQLATAS RSEEKDECAMISTSIGEEFELPISSATIKCAES LQPVAASVEREATGFVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTSALECEASVS GVVVESENERAGTVTEERGDGSGISTSSVEDC EGGPSSAVPQEEGDPSVTPAEEMGDTAMISTS TESEGCEAVMIGAVQAUERITTITVPELSDA AIISTSTABCMPISASIDRHEENQLTADNPEGN GGLEGPPVLAVSTEEGINGPSVIKKPSAGQGH PSAVCAEKEEKHIGKCPEIGPPAGRGESIST TESEGCEAVMIGAVLQUBENTITYRVEDLSDA AIISTSTABCMPISASIDRHEENQLTADNPEGN GGLEGPPVLAVSTEEGINGPSVIKKPSAGQGH PSAVCAEKEEKHIGKCPEIGPPAGRGESST	J	1]			
SRILCENRGGISLOGRAKGEEKLAANTLSTP SGSSLORPKKSODMTLPEOEPMEIDSEPGYE NVFEVSKTQDNRNNNSHQDIDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGGTGV NSNSEKHADHRISTLTKKMHIQSAVSKMNPGE KEPHRGTTEVNIDSETVHRMLSAPSINDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLBGSTA STSPADHSALPNQSLTVRESPVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVMPLGSKLTGVTVENENITKEGGLVDMA KKENDLNABEPNLAQTIKATVENGKKDGJAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEDVDTSAGSSAPSVLHQRNGQTE DVATOPRRABKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESERAADL LAVHAVKLENNVSVVTEEDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFTISEVESDGAVTSAG TEIRAGISSSEVDGSGGMMRMMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPSACTOSEDSSEGFAIS SSEEENGESAMDSTVAKEGTNVPLVAAGPCD DEGITVSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSGIGTVEH VEAAEGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTNQNSLAGGNQGK VLIJSTSTTNDTTPQVSATTDVPGGGLSDALRTE ENNEGTRVTTEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGGEFISSATIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAP EAESH ASTSKEEKDEGALISTSAAECEASVS GVVVESENERAGTWTHEEPSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAP EAESH ASTSKEEKDEGALISTSAAECEASVS GVVVESENERAGTWTHEEPSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAP EAESH ASTSKEERDGSGISTSSVEDC EGPVSSAVPQEEGPSVTTAEEMGDTAMISTS TESGCAVMIGAVLQDENITTITVEDISDA AIISTSTAECMPISASIDRHEENQLTAMISTS TESGCAEVINGGAVLQDENITTITRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMISTS TESGCAEVINGGAVLQDENITTITRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMISTICAES LQPVAAAVEERATGPVLISTADFEGYPAR GGKEPGPVLAVSTEEGHNGPSVHKPSAQGGH PSAVCAEKEEKHIGKRCFEIGPFAGROKESTL HLINAEENVILLISLOKGKOKSPETTATAGGSST ASYSAGRGLEGNANSPAHLRCPPGTFGQTAK DSSVSSINYLLAANNTGANSPAHLRCPPGTFGQTAK DSSVSSINYLLAANNTGANSPAHLRCPPGTFGQTAK		1					
SGSSLQRPKKSGDMTLIPEQEPMEDSEPGYE NVFEVSKTQDNRNNSHQDDSENMKQKTS ATVQKDELRICTADSKATAPAYKPGRGTGV NSNSEKHADHRSTITKKIMIIQSAVSKMPGE KEPIRIGTTEVNIDSETYEVHIMLISAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDPVIPDKITVLEGSTA STSPADHSALIPNQSLTVRESEVLKTJDSKEGG GEFTVDTPAKASITSKRHIPAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPILKQTIKATVENGKKDGLAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERNINSEVDITSAGSGSAPSVLHQNGQTE DVATGPRRAKTSVATSTEGKDKDVTLSPVK AGPATITTSETRQSEVALPCTSIEADGLIIGT HSRNNPLHVQAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIRANVASVVTEGESPRADL LAVHAVKIRANVASVVTEGNOPKKETEG TVTCTGAEGRSDNPVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDSVNDGTE GESAVTSTGITEDGGEPSCTTGSEDSSEGFAIS SESEENGESAMDSTVAKSGTNVFLVAAGPCD DEGTVTSTGAKEEDEGEGVTYSTGRGPREERMVT GAGVVLGDNDAPPGTSASQEGDSVNDGTE GESAVTSTGITEDGGEPSCTGSEDSSEGFAIS SESEENGESAMDSTVAKSGTNVFLVAAGPCD DEGTVTSTGAKEEDEEGGVTYSTGRGPREERMVT VAEAGAAMNANENNVDSMGTTEGSISCDT DICSSAKGTVSSSVTSAVSGREVEYPYPGGCE GPMTSAASDQSDSQLEKVEDTITSTSV NECCOGLMATTASGDTINQNSLAGGKNQGK VLUISTSTTINDTYPQVSATDYSGRGSLGSDALTE ENNEGTRVTTEEFEAPMPSAVSGDSOLTAS RSEEKDECAMISTSIGEEPELPISSATTIKCAES LQFVAAAVERSTATSVEKEDGGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAERMGDTAMISTS TESEGCAVMIGGAVGAVGREPSAPP EAESPLASTSKEEKDECALISTISAECEASSV GVVVESENERAGTVTMEERMGDTAMISTS TESEGCAVMIGGAVGGRESTSVEDC EGPVSSAVPQEEGDPSVTPAERMGDTAMISTS TESEGCAVMIGGAVGGRESTSVEDC GGREPGPVLAVSTEEGINFORGNORSTED GREPGPVLAVSTEEGINGPSVIKKPSAGGGH PSAVCAEKEEKHGKGEGIGSTSSVEDC GGPSSAVPGEGOPSVTPAERMGDTAMISTS TESEGCAVMIGGAVCDERDTTTRVPLISDA ALISTSTADECMIPSASUBRIERNQLTADNPEGN GDLSATEVSKIKVPMPSLLAENNCRCPGPVR GGREPGPVLAVSTEEGINFORGNORSTIL HLINAEENNVLAANVIGGAUGRSTIL HLINAEENNVLAANVIGGAUGRSTIL HLINAEENNVLAANVIGGAUGRSTIL HLINAEENNVLAANVIGGAUGRSTICANSTICAGSST ASYSAGGGLEGNANSPAHLRCPEGTSGGTAK DSSVSSINYLAAANVIGGAUGRDMPPVOGTVA	1	1		1			
NVFEVSKTQDNRNNSHQDIDSENKQKTS ATVQKDELRICTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKAMPGE KEPHRGTTEVNIDSETVHRMLISAJSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPLFPKRTVLBGSTA STSPADHSALPNQSLTVRESVLKTDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENGKKDGIAVD HVVGLNTEKYAETVKLKHRRSPGKVKDISID VERRIENSEVDTSAGGSSVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTISSETTGSEVALFSTEABEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFASSETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFTISEVESDGAVTSAG TERAGSISSEVDGSSGOMMRMMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSAGGGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSGJCTVVEH VEAAEGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDOSDSOLEKPTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDTVTPQVSAGTTISTGATKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESGENERAGTVTMEERJGSGTISTSSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMISTT TESGCCAVMIGAVLQDEDTTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESGENERAGTVTMEERGOGGISTSSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMIST TESGCCAVMIGAVLQDEDTTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESGENERAGTVTMEERGOGGISTSSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMIST TESGCCAVMIGAVLQDENTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP EAESPLASTREEKGDRIGGTSTSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMIST TESGCCAVMIGAVLQDENTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP GGKEPGPVLAVSTEEGHNGPSVHKPSAQGGH PSAVCAEKEEKHGKRCPEIGPFAGROKESTL HLINAEENVALLANNTGGLEGDPTYV	}	1 .		1			
NVFEVSKTQDNRNNSHQDIDSENKQKTS ATVQKDELRICTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKAMPGE KEPHRGTTEVNIDSETVHRMLISAJSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPLFPKRTVLBGSTA STSPADHSALPNQSLTVRESVLKTDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENEVIKTDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENGKKDGIAVD HVVGLNTEKYAETVKLKHRRSPGKVKDISID VERRIENSEVDTSAGGSSVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTISSETTGSEVALFSTEABEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFASSETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFTISEVESDGAVTSAG TERAGSISSEVDGSSGOMMRMMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSAGGGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSGJCTVVEH VEAAEGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDOSDSOLEKPTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDTVTPQVSAGTTISTGATKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESGENERAGTVTMEERJGSGTISTSSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMISTT TESGCCAVMIGAVLQDEDTTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESGENERAGTVTMEERGOGGISTSSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMIST TESGCCAVMIGAVLQDEDTTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESGENERAGTVTMEERGOGGISTSSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMIST TESGCCAVMIGAVLQDENTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP EAESPLASTREEKGDRIGGTSTSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMIST TESGCCAVMIGAVLQDENTTTRVEDISDA AIISTSTAECMPISASIDRHEENQLTAMSPEAP GGKEPGPVLAVSTEEGHNGPSVHKPSAQGGH PSAVCAEKEEKHGKRCPEIGPFAGROKESTL HLINAEENVALLANNTGGLEGDPTYV		'					SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE
ATVOKDELRTCTADSKATAPAYRFORGTGV NSNSEKLADHRSTI-KAMHOGA AYSKAMPOG KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV QKNILKNTAAEEHVAQQDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGGTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG GEGTVDTPTAKASITSKHPBAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENEKKDGGAVD HVVGLNTEKYAEFTVKIKKRSPGKVKDISID VERRNENSEVDTSAGGSAAPSVLHORNGQTE DVATOPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGALASECTVFAAAEEGGAVVTE GFASESTFLTSTKEGESGECAVABSEBRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSSDSETVEGTILTSEVESGGAVTSAG TEIRAGSISSEVDGSQGNMMEMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTTEDGEOPASCTGSEDSSEGFAIS SESEENGESAMDSTTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTSTGRGNEIGH ASTCTGLGEESGGVLICESABGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GFMTSAASDQSBOGLEVEDTTSTGRJVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLAMTTASGDTNONSLAGGKNOGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEKENGECAMISTSTIGEFFELPISSATTIKCAES LQPVAAAVEERAATGPVLISTADFGOPMPSAPP EASPALSATSKEEKDECALISTSIAEPGCASVS GVVVESENERAGTVMEKDGSGISTSSVEDC EGFVSSAVRQEEGDBSVTTAEEMGTAAMSTS SESEENGECAMSTSTIGEFFELPISSATTIKCAES LQPVAAAVEERAATGPVLISTADFGOPMPSAPP EASPALSTSKEEKDECALISTSIAEPGCASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGFVSSAVRQEEGDBSVTTAEEMGTAAMSTS TSEGCEAVMGAVLQDERLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEGGN GDLSATEVSKHV WPDGARGGTAKSTS TSEGCEAVMGAVLQDERLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEGGN GDLSATEVSKHV WPDGARGGTAKSTS LASVAGRGLEGNANSPAHLRGPEGTSGGTAK DSSVSISTRLAAVNTGAKEDMSPVGTVA	}	1		j			
NSNSEKHADHRSTLTKKMHIQSA VSKMAPGE KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV QKNILKNTAAEEHVAQGDATLEHSTINLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGSTA STSPADHSALPNQSLTRSEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGJAVD HVVGLNTEKYAETVALKHKRSPGKVKDISID VERRNENSEVDTSAGSAPSVLHQRNGQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVISVETEEDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFISEVESDGAVTSAG TEIRAGSISSEVVDGSQONMMRMGYKKETEG TVTCTGAEGRSDNFVCSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDOTE GESAVTSTGITEDGEGDSTSCTSGEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTSTGRGNEIGH ASTCTGLGEESSCVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMGTEGGSKDT DICSSAKGVESSVTSAVSKCBEVTPYPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVECEVAHTSPSEKEDEDIITTSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLIISTSTTNDYTTPQVSAITDVEGGLSDALRTE ENMEGTRYTTEFEFAMPSAVSGDDSQLTAS RSEKEDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFGPMPSAPP EASPLASTISKEEKDECALISTISIAEECEASVS GVVVESENERAGTVMEKKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TEBGCEAVMIGAVLQEDRLTTTRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNESDN GGLSATTSKEKPLOEALISTISIAEECEASVS GVVVESENERAGTVMEKKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TEBGCEAVMIGAVLQEDRLTTTRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNESDN GDLSATEVSKHVKPMFAEEMGDTAMISTS TEBGCEAVMIGAVLQEDRLTTTRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNESDN GGLSATEVSKHVKPMISLENDNCCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKRPSAGOGH PSAVCAEKEEKHGKECPEIPFAGRGQKESTIL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTSGQTAK DSSVSSINVLAAVNITGAKADDMPYOGTVA		1					ATVOKDELRTCTADSKATAPAYKPGRGTGV
KEPIHRGTTEVNIDSETVHRMLISAPSENDRV OKNI,KNTAAEEHVA,OATLEHISTNILDSSPS LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASTISKRKHIPEAHQATLLDGKQ GKVIMPLGSKI,TGVIVENENITKEGGLVDMA KKENDLNAEPHLK,OTKATVENGKKDGJAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVHLPCNSTGENE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVDT GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEVDGSQONMMRGPKKETEG TVTCTGAEGRSDNPVCSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQECDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTSTGRGNEIGH ASTCTGLGESSGOVLGESAEGDSQIGTVVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSKDEVTPVPGGCE GPMTSAASDQSDQSLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDTINONSLAGGKNOGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTTEFEAPMPSAVSGDDSQLTAS RSEKNDECAMISTSIGEFEFLPISSATTIKCAES LQPVAAAVEERATGPVLISTADFGPMPSAPP EASPLASTISKEENDECALISTSIAEEGEASVS GVVVESENERAGTVMEKEDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATTSVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATTSVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATTSVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATTSVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATEVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATEVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATEVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATEVSKHV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATEVSKHKV PROEDRLTTMRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNEEGN GDLSATEVSKHKV PROEDRLTTMRVEDLSDA AIISTSTAECKPISASIDRHEENQLTADNEEGN GDLSATEVSKHKV PROEDRLTTMRVEDLSDA AIISTSTAECKPISASIDRHENQLTADNEEGN GDLSATEVSKHKV PROEDRLTTMRVEDLSDA AIISTSTAECKPISASIDRHENQLTADNEEGN GDLSATEVSKHKV PROEDRLTTMRVEDLSDA AIISTSTAECKPISASITAGENCECTOFTV		1					
QKNLKNTAAEHYAQGDATLEHSTNLDSSPS LSSVTVVPLRESYDPOVPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTFAKASITSKRIIFEAHQATLLDGKQ GKVIMPLGSKLT GVIVENENITKEGGLVDMA KKENDLNAEPILKQTIKATVENGKKDGLAVD HVVGLNTEKYAETVKLKKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTEGDKDVTLSPVK AGPATTTSSETRQSAUPCTSSEADGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFASSETFLTSTKEGESGECAVAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSE KCDGSLSRDSEVVEGTITTISEVSDGAVTSAG TEIRAGSISSEEVDGSOMMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPRGTSASQEDGSVNDOTE GESAVTSTGITEDGEGPASCTGSEDSEGFAIS SESSENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVYTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPYPGGCE GPMTSAASDQSDQLEKVEDTTISTELVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTTNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRYTTEEFEAPMPSAVSGDDSQLLARTE ENMEGTRYTTEEFEAPMPSAVSGDDSQLARTE ENMEGTRYTTEEFEAPMPSAVSGDSQLARTE ENMEGTRATTVATHERENCLATADTEGN GGKEFGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAEKEEKHGKEPCFIGPFAGGGKSTL HILINAEEKNVLLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGFGQTSGGTAK DSSVSSSRYLAAAVNGAIKARDDMPVQGTVA	1	1 .)			
STSYTVYPLRESYDPDVIPLFDKRITVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGITVDTTAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEFNILKGTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNPVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDOTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENVDSMMSTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDUITSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFEFLISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDTYTPAEMGGTAMISTS TSEGCEAVMIGAVLQDEDRLTTIR VEDLSDA AIISTSTAECMPISASIDRHEENOLTADNPEGN GDLSATEVSKIKKVPMPSLLANDRCCPGVR GGKEFGPVLAVSTEEGHNORSVHKPSAGQGH PSAVCAKEEKHRIGSCPFIGFFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGFEGTSGGTAK DSSVSSIRYLAAAVNIGAIKADDMPPOGTGAK DSSVSSIRYLAAAVNIGAIKADDMPPOGTGAK DSSVSSIRYLAAAVNIGAIKADDMPPOGTGAK DSSVSSIRYLAAAVNIGAIKADDMPPOGTGAK DSSVSSIRYLAAAVNIGAIKADDMPPOGTGAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK DSSVSSIRYLAAAVNIGAIKADDMEPONGTTAK				l			
STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVGLNTTEKYAETVKHKKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATOPRAEKTSVATSTEGOKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADGGLIGT HISRNNPLHVGAEASECTVAAAEEGGAVVTE GGAESETFLTSTKEGGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSADSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESSENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTTGAKEEDGEDVVTSTGRONEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKOT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVASTSPSEKEDEDITTSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVFGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAK RSEEKDECAMISTSIGGEFELPISSATTIKCAES LQPVAAAVSERAFGVSTSTAVSGKOEVTYPAGGE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGGEFELPISSATTIKCAES LQPVAAAVSERAFGVSVTYAEGGGSGISTSSVEDC EGGVSSAVPQEEGDTYTAEEMGGTAMISTS SESELDECAMISTSIGGEFELPISSATTIKCAES LQPVAASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGGVSSAVPQEEGDTYTAEEMGGTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRIEEEMQLTADNPEGN GDLSATTASETROKTPROPTYPAGGGC GGKEFGPVLAVSTEEGHNGPSVHKPSAQGGH PSAVCAEKEEKHGKEPGIFFAGGGSST ASYSAGRGLEGNANSPAHLRGFEQTSGGTAK DSSVSSIRVLAAVNIGAIKADDMEPVGGTSA	1	i		l i			1
EGFTVDTPAKASITSKRHIPFAHQATLLDCKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPILKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKILKHRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEKDDAVTSAGSEE KCDGSLSRDSEIVGTTTFISEVESDGAVTSAG TEIRAGSISSEVDGSQOMMRRMGPKKETEG TVTCTGAEGRSDNPVICSVTGAGPREERMVT GAGVVLGINDAPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDHITSVE NEECDGLMATTASGDITINQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFLPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEERDGGISTSVEDC EGPVSSAVPREGOPSVTPAEPGMTANISTS TSEGGEAVMIGAVLQDEDRITTITRVEDLSDA AIISTSTAECMPISAGTHERNGCTTARPEGN GDLSATESVELVERAGTHERNGTANISTS TSEGGEAVMIGAVLQDEDRITTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGGGTAK HUNAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTSGGTAK		[
GKVIMPLOSKL TGVIVENENITKEGGLVDMÅ KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRINPLHVGAEASECTVFAAAEEGGAVTE GFAESETFLTSKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQQOMMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTTEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTTGROEIGH ASTCTGIGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLIJSTSTNDYTPQVSAITDVEGGLSDALRIE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGFMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGITSSVEDC EGPVSSAVPQEEGDPSVTTAEEMGDTAMISTS TSEGCEAVMIGAVLQBERLITTITRVEDLSDA AIISTSTAECMPISASDRHEENDLTANDFEG GGKETGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAEKEEKHGCPEGIPFAGRGGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHRSGPGTSGTAK	į.						
KKENDLNAEPNILKOTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVOTSAGSGSAPSVLHQRNOQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVAAAAEEGGAVVTE GFAESETFILTSTKEGESGECAVAESEDRAADL LAVHAVKLEANNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGGPSSAVPQEGOPSVTTAEEMGDTAMISTS TSEGCEAVMGAVLQDEDRLTITRVEDLSDA ALISTSTADSTYMEEKDGSGIISTSVEDC EGGPSSAVPQEEGPPSVTTAEEMGDTAMISTS TSEGCEAVMGAVLQDEDRLTITRVEDLSDA ALISTSTAECHSASIDRHEENQLTADNPEGN GGKEPGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAKEKEKHGKECPEIGPFAGRGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPPGTSGCTAK DSSVSSIRYLAAAVNTGAKKADDMPPYOGTVA	ļ	1		,			EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ
HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRAÆTSVATSTEGKDKDVTLSPVK AGPATTTSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFITSTKEGESGECAVAESEDRAADL LAVHAVKIEANNNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKFTEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTTEDGEGFASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMGSTEKGSKDT DICSSAKGIVESVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATASGDTINQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENNEGTRYTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFEFLISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVSESENERGTVMEEKGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMGAVLQDDRLTTTRVEDLSDA AIISTSTACAPPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNQFSVHRPSAQOGH PSAVCAKEKERGHGRENGPSHKPSAQOGH PSAVCAKEKERGHGRENGPSHKPSAQOGH PSAVCAKEKERGHGREPGPTAGRGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGCTAK DSSVSSIRYLAAAVNTGAKKADDMPPYQGTVA							GKVIMPLGSKLTGVIVENENITKEGGLVDMA
VERRNENSEVDTSAGSGSAFSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGFATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETIFLTSTKEGESGECAVAESEDRAADL LAVHAVKEANNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGFASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRONEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGFMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEHNGPSVYHKPSAQGH PSAVCAEKEEKHGKECPEIGFPAGRGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPPQTSQTAK DSSVSSIRYLAAVNTGAKKADDMPPYOGTVA	1						KKENDLNAEPNLKQTIKATVENGKKDGIAVD
DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGMMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEINGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAMPSAVSGDDSQLARTE ENMEGTRVTTEEFEAMPSAVSGDDSQLAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTISAECCEASVS GVVVESENFAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITTRVEDLSDA AIISTSTAECMPISASIDRHEINTLTRVEDLSDA AIISTSTAECMPISASIDRHEINQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGRINGPSVHKPSAGQGH PSAVCAEKEKHGKCPEIGPFAGRGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGTAK DSSVSSIRVLAANNSPAHLRGPEQTSGTAK DSSVSSIRVLAANNSPAHLRGPEQTSGTAK	ſ	1					HVVGLNTEKYAETVKLKHKRSPGKVKDISID
DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGMMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEINGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAMPSAVSGDDSQLARTE ENMEGTRVTTEEFEAMPSAVSGDDSQLAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTISAECCEASVS GVVVESENFAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITTRVEDLSDA AIISTSTAECMPISASIDRHEINTLTRVEDLSDA AIISTSTAECMPISASIDRHEINQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGRINGPSVHKPSAGQGH PSAVCAEKEKHGKCPEIGPFAGRGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGTAK DSSVSSIRVLAANNSPAHLRGPEQTSGTAK DSSVSSIRVLAANNSPAHLRGPEQTSGTAK	1	1		[VERRNENSEVDTSAGSGSAPSVLHORNGOTE
AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESEGCAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGISSEVDGSQGMMRMGPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLUSTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECLASTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLINSLQKEDKSPSTTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTSGCTAK DSSVSSIRVLAAVNTGAIKADDMPPVOGTVA	1						
HSRNNPLHVGAAASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGINDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEDDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPQPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1	!		1	!		
GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGIGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEMGGTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAKKADDMPPVOGTVA							· ·
LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRGPKEFEG TVTCTGAEGRSDNIPVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGROREIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRIE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLITTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKIKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNIGPSVHRPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPTGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		}			-		
KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGROREIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTIRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	i	1 1			Í		
TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					ļ		1
TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENNEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGFMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGFSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	}	1 1				İ	
GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECGLISTSIAECCEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA		Ì					
GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSLAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1						TVTCTGAEGRSDNFVICSVTGAGPREERMVT
SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKCPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1	l i					GAGVVLGDNDAPPGTSASQEGDGSVNDGTE
SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKCPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1	j .				i	GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS
DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNYDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTIRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPTGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1	, ,			}	ļ	
ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1				l	
VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		'				ļ	I
DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1 1				i	
GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA]			.	l	
YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1) .	j	J	j	j	- 1
NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTI HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	j J	J	ı j		}	
VLIISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		j 1	- 1	1		j	
ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1			1			
RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA		į (ľ	l	1	ľ	
LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA		1 l	ļ	į			ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS
LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA		j j	- 1	j			RSEEKDECAMISTSIGEEFELPISSATTIKCAES
EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA		[1			1
GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		į į	1	l	1		,
EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1 1	1	i	i		
TSEGCEAVMIGAVLQDEDRLITTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		l l	l	l		,	
AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA]) l	1	ŀ]		
GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1]	ſ	ļ	1	ſ	
GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		1	ł			` '
PSAVCAEKEEKHGKECPEIGPFAGRGQKESTI, HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1	j 1	1	ł	1	ł	
HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1	1	.		1		
HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	1	ļ l			1		PSAVCAEKEEKHGKECPEIGPFAGRGOKESTL
ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA	i	į į	1]	ĺ	1	
DSSVSSIRYLAAVNTGAIKADDMPPVOGTVA		[[Ì	- 1	ľ	1	
EHSFLPAEQQGSEDNLKTSTTKCTTGQESKIAP						i	
	1	}	})	ļ	.]	ERCEL DY EUUCGEDVIL KAGAMACAMACAMACAMACAMACAMACAMACAMACAMACAM
	L	LL	i		l	1	ENSILPAEQQUSEUNLK 1511 KCITGQESKIAP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, =possible nucleotide deletion, \=possible nucleotide insertion SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK
						WTDQASAEKTGDDNSTRKSFPEEGDIMVTVS SEENVCDIGNEESPLNVLGGLKLKANLKMEA YVPSEEEKNGEILAPPESLCGGKPSGIAELQRE PLLVNESLNVENSGFRTNEEIHSESYNKGEISS GRKDNAEAISGHSVEADPKEVEEEERHMPKR KRKQHYLSSEDEPDDNPDVLDSRIETAQRQC PETEPHATKEENSRDLEELPKTSSETNSTTSRV MEEKDEYSSSETTGEKPEQNDDDTIKSQE
1302	2652	A	10167	321	842	EPSLFPFLRPSPARPPPRPPAPFPSPELAGPEPH FVFYFFLSYVHPPKELAKYEYMEEQVILTEKG NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP EARETKCYVRSSVGCVEPLTTQAEVTENLDR KNSQQVFKLLKKK
1303	2653	A	10171	206	429	NMILLKKRRLLINSLGEGTINGLLDELLETNV LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA RACEICITYI
1304	2654	A	10184	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF YHEAVVLFTQALKLNPQDHRLFGNRSFCHER LGQPAWALADAQVALTLRPGWPRGLFRLGK ALMGLQRFREAAAVFQETLRGGSQPDAAREL RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLLH
1305	2655	A	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ SDFLTPPVGGAPWAVATTVVMYPPPPPPHR DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRNMILFLLAFLLFCGLLFYINLADHWKG IRNTCT
1306	2656	A	10195	1	410	IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTFTITVDQKTFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK
1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR
1308	2658	A	10214	2	453	ECGGIRQPGPGPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM
1309	2659	A	10233	45	421	RGWPEQOSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV
1310	2660	A	10241	243	442	AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK
1311	2661	A	10261	751	176	LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG
1312	2662	A	10270	3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP SMTILDKKDGEQAKALFEKVRKFRAHVEDSD

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ	Predicted	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	noa	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	согтеspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	1		peptide	}	/=possible nucleotide deletion, \=possible
ļ -	<u> </u>	-		sequence		nucleotide insertion LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF
ł	}	1		l	Į	EHVCKPKVEHLIGYEVFECTHNMAYMLKKL
Į.			1			LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV
İ	{		i		Ì	REESSFSDIPDVKNDFAFLLHMVDQYDQLYS
			1			KRFGVFLSEVSENKLREISLNHEWTFEKL
1313	2663	A	10287	1221	266	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR
1						VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL
	}		1			LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE
1						EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD
1		1				GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL
		Ì				KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW
1						GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF
1						PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL
1						LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL
1				•		VKKDKSEKELKALCIDQLDVFLQKETQIFVEK
	i					LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK
		l				DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR
						ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD
1		l				RYNRRRGRSRSYSRSRSRSWSKERLRERDRD
		l				RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN
· [1				NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG NNTTESWSEFHEDQVDHNSYVRPPMPKKRC
						RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN
					·	LPGMQPFPAQPPVVEGPPPPGLPPPPPILTPPPV
						NLRPPVPPPGPLPPSLPPVTGPPPPLPPLQPSG
						MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT
1		[ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR
1315	0666	<u> </u>	10000		1001	AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAEERRTRGA
1						GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI GTPSPCGSAQAAAAAAAEEATEKIPALRPALL
						WALLALWLCCATPAHALQCRDGYEPCVNEG
						MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE
1	1					KNRCQNGGTCVAQAMLGKATCRCASGFTGE
						DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE
						CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS
						GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT
10.5	2411		1000			LWVMVSGATSTSTSTL
1316	2666	A	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG
						YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI
		'				GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR
						NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	Α	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPO
/	200,	''			1550	DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP
						LPAASSGMKSSKSSTSLAFESRLSRLKRASSE
						DTLNKPGSTAASGVVRLKKTATAGAISELTES
				1		RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV
					İ	PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP
		ĺ		[ĺ	KQENEGGEKAALESQVRELLAEAKAKDSEIN
]]					ļ	RLRSELKKYKEKRTLNAEGTDALGPNVDGTS
						VSPGDTEPMIRALEEKNKNFQKELSDLEEENR
			İ	ļ		VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ
				J	}	ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS
						NPFKSSKCSTAGSSPNSVSELSLASLTEKIOKM
1	i		ļ		ļ	EENHHSTAEELQATLQELSDQQQMVQELTAE
<u> </u>						

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	100	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	1	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uu,ice	l	914	correspondi ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
den		1	717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
1		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1		peptide	sequence	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
}	ļ	<u> </u>		Sequence	 	NEKLVDEKTILETSFHQHRERAEQLSQENEKL
j		j	J		[MNLLQERVKNEEPTTQEGKIIELEQKCTGILE
l .			1		ļ	
					ł	QGRFEREKLLNIQQQLTCSLRKVEEENQGAL EMIKRLKEENEKLNEFLELERHNNNMMAKTL
1		1			i	EECRYTLEGLKMENGSLKSHLOG
1318	2668	A	10303	333	879	
1310	2006	A	10303	333	6/9	GECFIMAAVVQQNDLVFEFASNVMEDERQL
		ì				GDPAIFPAVIVEHVPGADILNSYAGLACVEEP
1						NDMITESSLDVAEEEIIDDDDDDDTTLTVEASCH
1		1	[ĺ	DGDETIETIEAAEALLNMDSPGPMLDEKRINN
}	ľ	1	1			NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ
1210	2660	<u> </u>	10222	160	754	QVQEKYADSPGASSPEQPKRKKK
1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLLIIATALSL
		1				LIGAKSLPASVVLEAFSGTCQSADCTTVLDAR
		1	j .		ļ	LPRTLAGLLAGGALGLAGALMQTLTRNPLAD
ļ		l				PGLLGVNAGASFAIVLGAALFGYSSAQEQLA
}		1]	MAFAGALVASLIVAFTGSQGGGQLSPVRLTL
1000	0.000	 	10000		<u> </u>	AGVXL
1320	2670	A	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV
1		ł			<u>.</u>	AVVDIQSDKAANVAQEINAEYGESMAYGFG
						ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI
1		ļ] ;			AKAAFISDFQLGDFDRSLQVNLVGYFLCARE
						FSRLMIRDGIQGRIIQINSKSDE
1321	2671	Α	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY
			1			AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD
		ĺ	[ILKCTLLVFGVRILYILKLNYTTEECDMKNMH
1						YVDPDHVKRAQKYAQQVLQKESPPKFAKTS
						MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	Α	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS
)			,			ERKMRAHQVLTFLLLFVTTSGASENASTSRGC
		1	1			GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA
			i			LITLLLMLILLGRLPFIKEKEKKSPAVLHFLFL
Ĺ						LGTLG
1323	2673	Α	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH
						QVLTFLLLFVITSVASENASTSRGCGLDLLPQ
'						YVSLCDLDAIWGIVVEAAAGAGALITLLLMLI
			l			LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324	2674	Α	10336	1	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE
		, .				NSVTHHEVKCQGKPLAGIYRKREEKRNAGN
						AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA
						AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA
						PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE
]]						LQSEERKRIDELIESGKEEGMKIDLIDGKGRG
[VIATKQFSRGDFVVEYHGDLIEITDAKKREAL
	1					YAQDPSTGCYMYYFQYLSKTYCVDATRETN
}					·	RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS
L				[RDIAAGEELLYDYGDRSKASIEAHPWLKH
1325	2675	Α	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL
			l	i		RGVTATFGRPAEWPGYLSHLCGRSAAMDLG
	ļ					PMRKSYRGDREAFEETHLTSLDPVKOFAAWF
})					EEAVQCPDIGEANAMCLATCTRDGKPSARML
1 1				ŀ		LLKGFGKDGFRFFTNFESRKGKELDSNPFASL
j l				l		VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS
) [[j	ĺ	RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE
					ĺ	QLYQDQEVPKPKSWGGYVLYPQVMEFWQG
				ļ	. 1	QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE
	ł					DWLYERLAP
1326	2676	Α	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV
				•		LLASLGVGLVTLLGLAVGSYLVRRSRRPOVT
						LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA
						The state of the s

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i]		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł	Į.	j		peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
		<u> </u>	 	sequence	ļ	HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD
			<u> </u>]		EDQGYVDLVIKVYLKGVHPKFPEGGKMSQY LDSLKVGDVVEFRGPSGLLTYTGKGHFNIOP
)				NKKSPPEPRVAKKLGMIAGGTGITPMLOLIRA
1	Ì					ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ
,]	}	İ	ļ		ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD
		ĺ	ļ			MIREHLPAPGDDVLVLLCGPPPMVQLACHPN
		1	j			LDKLGYSQKMRFTY
1327	2677	Α	10345	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRR
						RYHRLSDMSMLAERRRKQKWAVDPQNTAW
ļ		ļ	ļ			SNDDSKFGQRMLEKMGWSKGKGLGAQEQG
		j				ATDHIKVQVKNNHLGLGATINNEDNWIAHQ
ļ]	J			DDFNQLLAELNTCHGQETTDSSDKKEKKSFS
			1			LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL
						DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF TIQEYFAKRMAALKNKPQVPVPGSDISETQVE
		-				RKRGKKRNKEATGKDVESYLQPKAKRHTEG
		ŀ				KPERAEAQERVAKKKSAPAEEQLRGPCWDQ
			1			SSKASAQDAGDHVQPA
1328	2678	A	10346	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGI
						CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF
						HMHCILKWLHAQQVQQHCPMCRQEWKFKE
1329	2679	Α	10351	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL
						SMYLVTVLGNLLILATISDSHLHTPMYFFLSN
						LSFADICVTSTTIPKMLMNIQTQNKVITYIACL
						MQMYFFILFAGFENFLLSVMAYDRFVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLOI
!						LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF
						LNHMVIYFTVALLGGGPLTGILYSYSKIISSIH
i		1				AISSAQGKYKAFSTCASHLSVVSLFYGAILGV
						YLSSAATRNSHSSATASVMYTVVTPMLNPFI
						YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680	Α	10352	34	2573	IPFLKSCCCCCLFDFPPPPLDQVQEFECEVERV
						TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE
						TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL
						FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS
						VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNO
ł (1		l	PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR
ĺ			1		ļ	LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT
1					i	EREKVKKAADHCRQILNYVNQAVKEAENKQ
						RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK
}				i	1	RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV
						LLQKQDDRLVLRCHSKILASTADSKHTFSPVI
} }						KLSTVLVRQVATDNKALFVISMSDNGAQIYE
						LVAQTVSEKTVWQDLICRMAASVKEQSTKPI
} }				ŀ		PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD
	1					LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS
	1			-		HLPVSEERWALDALRNLGLLKQLLVQQLGLT
	ľ	. 1		İ		EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE
	ļ					NIKAYHSGEGHMPFRTGTGDLATCYSPRTSTE
	ľ					SFAPRDSVGLAPQDSQASNILVMDHMIMTPE
						MPTMEPEGGLDDSGEHFFDAREAHSDENPSE
	1	- 1			l	GDGAVNKEEKDVNLRISGNYLILDGYDPVQE
					ļ	SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP
		l	1		1	QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI
]			QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE SYTILCORLAGSALTDKHSDKS
i					1	911ILCAVEWOOVE INVUSING

1 7 1	QID Met	SEQ	Predicted	Predicted end	
1 1 -	of hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide seq	otide	USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq- uer	•	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
uence		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
acinoc	l	7,14	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1 1	- {		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 1	- 1	1	peptide) sequence	/=possible nucleotide deletion, \=possible
	- 1		sequence		nucleotide insertion
1331 268	31 A	10353	1	2100	AVEFAEGALTMAPWPELGDAQPNPDKYLEG
	1		1]	AAGQQPTAPDKSKETNKTDNTEAPVTKIELLP
	1	Ì			SYSTATLIDEPTEVDDPWNLPTLQDSGIKWSE
			İ		RDTKGKILCFFQGIGRLILLLGFLYFFVCSLDIL
1 1	1	1	ł	Ì	SSAFQLVGGKMAGQFFSNSSIMSNPLLGLVIG
1 1		1	ł		VLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
1	1		•		IIMGANIGTSITNTIVALMQVGDRSEFRRAFA
	ļ				GATVHDFFNWLSVLVLLPVEVATHYLEIITQL
l i	1				IVESFHFKNGEDAPDLLKVITKPFTKLIVQLDK
	ŀ				KVISQIAMNDEKAKNKSLVKIWCKTFTNKTQ
	[INVTVPSTANCTSPSLCWTDGIQNWTMKNVT
	ľ				YKENIAKCQHIFVNFHLPDLAVGTILLILSLLV
1	ļ	1			LCGCLIMIVKILGSVLKGQVATVIKKTINTDFP
		}			FPFAWLTGYLAILVGAGMTFIVQSSSVFTSAL
		1			TPLIGIGVITIERAYPLTLGSNIGTTTTAILAAL
					ASPGNALRSSLQIALCHFFFNISGILLWYPIPFT
1 1					RLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIP
					LTVFGLSLAGWRVLVGVGVPVVFIIILVLCLR LLQSRCPRVLPKKLQNWNFLPLWMRSLKPW
					DAVVSKFTGCFOMRCCCCCRVCCRACCLLC
1	l		·		GCPKCCRCSKCCEDLEEAQEGQDVPVKAPET
1 1	1				FDNITISREAQGEVPASDSKTECTAL
1332 268	2 A	10354	30	1377	SQQGSQPHRQGPPSLLTAPHSLDLPALPPGPR
				-5,,	GSQGKLRRVLVPMSVKPSWGPGPSEGVTAVP
1 1	1				TSDLGEIHNWTELLDLFNHTLSECHVELSQST
i i	ĺ				KRVVLFALYLAMFVVGLVENLLVICVNWRG
		1			SGRAGLMNLYILNMAIADLGIVLSLPVWMLE
1	1				VTLDYTWLWGSFSCRFTHYFYFVNMYSSIFF
]]	ļ			ļ	LVCLSVDRYVTLTSASPSWQRYQHRVRRAM
					CAGIWVLSAIIPLPEVVHIQLVEGPEPMCLFM
, ,					APFETYSTWALAVALSTTILGFLLPFPLITVFN
1 1		1			VLTACRLRQPGQPKSRRHCLLLCAYVAVFV
1 .	1	1			MCWLPYHVTLLLLTLHGTHISLHCHLVHLLY
	ŀ				FFYDVIDCFSMLHCVINPILYNFLSPHFRGRLL
]	ļ				NAVVHYLPKDQTKAGTCASSSSCSTQHSIIIT
]	J		İ		KGDSQPAAAAPHPEPSLSFQAHHLLPNTSPISP
1333 268	3 A	10358	2	884	TQPLTPS
208.	, A	10226	4	004	AAGAGADGREPASERASRAEPPAVAMGQND
1	1	1 1	ł		LMGTAEDFADQFLRVTKQYLPHVARLCLIST
1 1	Ì	1 1	1	1	FLEDGIRMWFQWSEQRDYIDTTWNCGYLLA SSFVFLNLLGQLTGCVLVLSRNFVQYACFGLF
					GIIALQTIAYSILWDLKFLMRNLALGGGLLLL
	ļ]	ļ	ļ	LAESRSEGKSMFAGVPTMRESSPKQYMQLGG
]]	}	1	RVLLVLMFMTLLHFDASFFSIVQNIVGTALMI
[[[ĺ	LVAIGFKTKLAALTLVVWLFAINVYFNAFWT
			ļ		IPVYKPMHDFLKYDFFQTMSVIGGLLLVVAL
<u> </u>		1	_	- 1	GPGGVSMDEKKKEW
1334 2684	4 A	10367	59	1562	QAWSLQVALSPFFFPASPSNSFAAAVPQLLFP
				Ì	ELPLPHVPGQESAKRRSARRFLIMSELTKELM
1 1	ļ] }	ŀ		ELVWGTKSSPGLSDTIFCRWTQGFVFSESEGS
[[Ţ	ALEQFEGGPCAVIAPVQAFLLKKLLFSSEKSS
[[1 1]	ľ	WRDCSQEEQKELLCHTLCDILESACCDHSGS
į į	1	1 1	i	i	YCLVSWLRGKTTEETASISGSPAESSCQVEHS
		ļ	1	ĺ	SALAVEELGFERFHALIQKRSFRSLPELKDAV
]	1	}	J	J	LDQYSMWGNKFGVLLFLYSVLLTKGIENIKN
	}	j j	ļ	ļ	EIEDASEPLIDPVYGHGSQSLINLLLTGHAVSN
			1		VWDGDRECSGMKLLGIHEQAAVGFLTLMEA
		}	j	Į.	
]					LRYCKVGSYLKISKIPYLDCLASETHLTVFFA

SEQ ID NO: of NO: of nucl- ocide seq- uence	dine, line, an, don, ble
cotide sequence Sequence USSN	line, an, ion, ble
sequence uence 09/496 correspondi ng to first amino acid residue of peptide sequence 1914	an, ion, ble
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 915 100 100 100 100 100 100 100	an, ion, ble
amino acid residue of peptide sequence T=Threonine, V=Valine, W=Tryptoph Y=Tyrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possil nucleotide insertion LGIILLGPFLQEFFPDQGSSGPESFT KQSNYNEKVMYVEGTAVVMGFE DTPIKRCLQTKWPYIELLWTTDRS: 1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	ion, ble
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possi nucleotide insertion LGIILLGPFLQEFFPDQGSSGPESFT KQSNYNEKVMYVEGTAVVMGFE DTPIKRCLQTKWPYIELLWTTDRS: 1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	ion, ble
peptide sequence /=possible nucleotide deletion, \=possi nucleotide insertion LGIILLGPFLQEFFPDQGSSGPESFT KQSNYNEKVMYVEGTAVVMGFE DTPIKRCLQTKWPYIELLWTTDRS 1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	ble
sequence nucleotide insertion LGIILLGPFLQEFFPDQGSSGPESFT KQSNYNEKVMYVEGTAVVMGFE DTPIKRCLQTKWPYIELLWTTDRSI 1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	
LGIILLGPFLQEFFPDQGSSGPESFT KQSNYNEKVMYVEGTAVVMGFE DTPIKRCLQTKWPYIELLWTTDRSI 1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	WINDLET
KQSNYNEKVMYVEGTAVVMGFE DTPIKRCLQTKWPYIELLWTTDRS 1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	
1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	
1335 2685 A 10375 82 2929 TRTKRRLGREKAMASPPRGWGCG LGTLCEPGSGQIRYSMPEELDKGSI	
LGTLCEPGSGQIRYSMPEELDKGSI	
	ALNPRSGS
LVTAGRIDREELCAQSPLCVVNFN	ILVENKM
KIYGVEVEIIDINDNFPRFRDEELKY	
AAGTRLVLPFARDADVGVNSLRSY	
FSLDVVSGTDGQKYPELVLEQPLD	
DLLLTALDGGDPVLSGTTHIRVTV	
PLFTPSEYSVSVPENIPVGTRLLML	
GINGKLTYSFRNEEEKISETFQLDS	
QSLDYEESRFYLMEVVAQDGGAL VTVQDVNDNAPEVILTSLTSSISED	1
LFSVHDGDSGENGEIACSIPRNLPFI	
NYYHLLTTRDLDREETSDYNITLT	
PLSTESHIPLKVADVNDNPPNFPQA	
ENNPRGVSIFSVTAHDPDSGDNAR	
DTFQGAPLSSYVSINSDTGVLYALI	
RDLQLWVTASDSGNPPLSSNVSLS	
DNTPEILYPALPTDGSTGVELAPRS	AEPGYLV
TKVVAVDKDSGQNAWLSYRLLKA	
VGLHTGEVRTARALLDRDALKQSI	LVVAVED
HGQPPLSATFTVTVAVADRIPDILA	
IDPEDLDLTLYLVVAVAAVSCVFL	
LRLRRWHKSRLLQAEGSRLAGVPA	
DGVRAFLQTYSHEVSLTADSRKSH	
ADTLLSEESCEKSEPLLMSDKVDAI	
QQAPPNTDWRFSQAQRPGTSGSQN WPNNQFDTEMLQAMILASASEAAI	
GAGTMGLSARYGPQFTLQHVLQG	
QNVYIPGSNATLTNAAGKRDGKAF	
KKKSGKKEKK	AGOITON
1336 2686 A 10379 1 557 RPRRQPSFSCRVLVLEDPPCFRFT	NSMNOEK
LAKLQAQVRIGGKGTARRKKKVV	•
DKKLQSSLKKLAVNNIAGIEEVNM	IKDDGTVI
HFNNPKVQASLSANTFAITGHAEAI	
GILSQLGADSLTSLRKLAEQFPRQV	LDSKAPK
PEDIDEEDDDVPDLVENFDEASKNI	
1337 2687 A 10380 1 1263 IPGSTISWSPAAARGLSVCRCCRLH	
FGDLPEPERSPRPAAGKEAQKGPLI	
SSTDSGSGGPLLFDDLPPASSGDSG	
MVKTEGKGAKRKTSEEEKNGSEEL KASSVEGLYGVVAERVGEREEMO	
KASSVIFGLKGYVAERKGEREEMQ DITEECRPPSSLITRVSYFAVFDGHG	
AAQNLHQNLIRKFPKGDVISVEKTV	
TFKHTDEEFLKQASSQKPAWKDGS	TATCVI A
VDNILYIANLGDSRAILCRYNEESQ	KHAALSI
SKEHNPTQYEERMRIQKAGGNVRD	
LEVSRSIGDGQYKRCGVTSVPDIRR	COLTPND
RFILLACDGLFKVFTPEEAVNFILSC	
TREGKSAADARYEAACNRLANKAY	
NVTVMVVRIGH	·
1338 2688 A 10385 3 589 GPSQSMAAGELEGGKPLSGLLNAL	AQDTFHG
YPGITEELLRSQLYPEVPPEEFRPFL	1
KSIASADMDFNQLEAFLTAQTKKQ	
AAVISKFWKSHKTKIRESLMNQSRV	
LSWRVDGKSQSRHSAQIHTPVAIIEI	LELGKYG
QESEFLCLEFDEVKVNQILKTLSEVI	FFRIRITIE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QPN
1339	2089	A	10386	5 0	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG KCVICDSYVRPCTLVRICDECNYGSYQGRCVI CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL GSSKTDLFYERKKYGFKKR
1340	2690		10388		3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK RPSWMVDNKRMRTASNFQWLLSTFILLYLM NQVNSQKKGAPHDLKCVTNNLQVWNCSWK APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL IPDTPEILNLSADPSTSTLYLKWNDRGSVFPHR SNVIWEIKVLRKESMELVKLVTHNTTLNGKD TLHHWSWASDMPLECAIHFVEIRCYDNLHFS GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA LVGPRATSYTLVESFSGKYVRLKRAEAPTNES YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE NSSYLVALDKLNPYTLYTFIRCSTETFWKW SKWSNKKQHLTTEASPSKGPDTWREWSSDG KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV KWEDIPVEELRGFLRGYLFYFGKGERDTSKM RVLESGRSDIKVKNITDISQKTLRIADLQGKTS YHLVLRAYTDGGVGPEKSMYVVTKENSVGL IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN RSDAKPENHVVESYCPPIIEEEIPNPAADETGG TAQVIYIDVQSMYQPQAKPEEEQENDPVGGA GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP ND
1341	2691	A	10392	1	5057	MLPPKHLSATKPKKSWAPNLYELDSDLTKEP DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL IKLRNLCLDWLQPETRTKEEIIELLVLEQYLTII PEKLKPWVRAKKPENCEKLVTLLENYKEMY QPEGESLHGVLVVSAGLRCPLGLSASTLLTW SGLDNSLSWAAVGMSCVLWDIELHHDFLGV ATKSVSTHAQGDAAQGLGGTTVRMWARDSN LATGVLLDDNNSDVTSDDDMTRNRRESSPPH SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FROGFRFNSTLVSRKRVLERKRRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS SLSSPSFTESQPIDFGAMPYVCDECGRSFSVIS EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted and	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
[ĺ	ĺ	1	peptide	[]	/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
			<u> </u>	 `		SOVGGKRFECKDCGETFNKSAALAEHRKIHA
1	•	ł	ł		1	RGYLVECKNOECEEAFMPSPTFSELOKIYGK
1					i	DKFYECRVCKETFLHSSALIEHQKIHFGDDKD
1			1	ĺ	1	NEREHERERERGETFRPSPALNEFQKMYG
į					i	KEKMYECKVCGETFLHSSSLKEHQKIHTRGN
1		i		ł	!	PFENKGKVCEETFIPGQSLKRRQKTYNKEKLC
						DFTDGRDAFMQSSELSEHQKIHSRKNLFEGR
1		ŀ				GYEKSVIHSGPFTESQKSHTITRPLESDEDEKA
	,	į	ļ		ļ	FTISSNPYENQKIPTKENVYEAKSYERSVIHSL
				1	ļ	ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN
1		ł	}	}	1	HQRVRAGGNTSEGREYSRSVIHSLVASKPPRS
		!				HNGNELVESNEKGESSIYISDLNDKRQKIPAR
		1		Ì		ENPCEGGSKNRNYEDSVIQSVFRAKPOKSVP
			1			GEGSGEFKKDGEFSVPSSNVREYOKARAKKK
		1	1		[YIEHRSNETSVIHSLPFGEQTFRPRGMLYECO
				j		ECGECFAHSSDLTEHOKIHDREKPSGSRNYE
						WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK
						DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE
						NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ
1						KDDPDDKIYECEDCGLGFVDLTDLTDHOKVH
		'				SRKCLVDSREYTHSVIHTHSISEYORDYTGEO
1		i				LYECPKCGESFIHSSFLFEHORIHEODOLYSM
						KGCDDGFIALLPMKPRRNRAAERNPALAGSA
1						IRCLLCGOGFIHSSALNEHMRLHREDDLLEOS
						QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
						ESFVNPAELADHVTVHKNEPYEYGSSYTHTS
1						1
						FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE
1						LHLEEEEEDEAAAAAAAAAQEVEANVHVPQ
			, ,			VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV
						EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE
1						QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE
i						GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC
						HECTETFTSSTAFSEHLKTHASMIIFEPANAFG
()			:			ECSGYIERASTSTGGANQADEKYFKCDVCGQ
1342	2602		10702		1250	LFNDHLSLARHQNTHTG
1342	2692	Α	10393	2	1350	GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA
						ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS
						APAVLVVAVAVVVVVVSAVAWAMANYIHV
		'				PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ
						HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN
	i	1				TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL
					·l	QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH
		}		ļ		VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL
))						WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI
				ļ	l	LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII
]]	,	ļ			ļ	YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE
		1	·		Ì	FAGVSDVDYSLYPDRELQSQWLRAYLEAYK
)	j	,	j	J		EFKGFGTEVTEKEVEILFIQVNQFALASHFFW
		- 1				GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM
						KPEVTALKVPE
1343	2693	Α	10394	102	839	PEAQTSAVLAREKGHLPTMRHEAPMQMASA
1	ľ	ĺ		ĺ		QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK
	ľ				•	TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN
	l	ì	ĺ			EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI
	1	ļ				LMYDITNEESFNAVQDWSTQIKTYSWDNAQ
		1		İ		VILVGNKCDMEDERVISTERGQHLGEQLGFE
1 1	ļ	1	ļ	ļ	ļ	FFETSAKDNINVKQTFERLVDIICDKMSESLET
		j		j		DPAITAAKQNTRLKETPPPPQPNCAC
1344	2694	A	10395	2	4136	DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS
<u> </u>					1	LQQRTPAEMSPVLHFYVRPSGHEGAASGHTR

				.		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	}	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ļ	İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	ļ	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
<u> </u>		<u> </u>		sequence		nucleotide insertion
		1				RKLQGKLPELQGVETELCYNVNWTAEALPSA
1		1	ĺ	([EETKKLMWLFGCPLLLDDVARESWLLPGSN
	1	1	j	•		DLLLEVGPRLNFSTPTSTNTVSVCRATGLGPV
1		l				DRVETTRRYRLSFAHPPSAEVEAIALATLHDR
1		İ	ļ	ļ	•	MTEQHFPHPIQSFSPESMPEPLNGPINILGEGR
İ						LALEKANQELGLALDSWDLDFYTKRFQELQR
ļ						NPSTVEAFDLAQSNSEHSRHWFFKGQLHVDG
						QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA
i		1	İ		1	IQGKEVRFLRPEDPTRPSRFQQQQGLRHVVFT
					1	AETHNFPTGVCPFSGATTGTGGRIRDVQCTG
	ĺ				[RGAHVVAGTAGYCFGNLHIPGYNLPWEDLSF
	1	!				QYPGNFARPLEVAIEASNGASDYGNKFGEPV
}	I	1				LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS
1	1	1				MEADHISKEAPEPGMEVVKVGGPVYRIGVGG
			}		İ	GAASSVQVQGDNTSDLDFGAVQRGDPEMEQ
1	l	l	ł			KMNRVIRACVEAPKGNPICSLHDQGAGGNG
1						NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW
1	ł		}		}	GAEYQESNALLLRSPNRDFLTHVSARERCPA
1						CFVGTITGDRRIVLVDDRECPVRRNGQGDAP
1						PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP
1						MLQPLALPPGLSVHQALERVLRLPAVASKRY
	1					LTNKVDRSVGGLVAQQQCVGPLQTPLADVA
] .]				VVALSHEELIGAATALGEQPVKSLLDPKVAA
ł						RLAVAEALTNLVFALVTDLRDVKCSGNWM
•						WAAKLPGEGAALADACEAMVAVMAALGVA
1						VDGGKDSLSMAARVGTETVRAPGSLVISAYA
1	ľ					VCPDITATVTPDLKHPEGRGHLLYVALSPGQ
						HRLGGTALAQCFSQLGEHPPDLDLPENLVRA
ĺ	1	i				FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM
Ĭ						AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV
1		ĺ				LEVQEPDLAQVLKRYRDAGLHCLELGHTGE
						AGPHAMVRVSVNGAVVLEEPVGELRALWEE
.]						TSFQLDRLQAEPRCVAEEERGLRERMGPSYC
<u> </u>						LPPTFPKASVPREPGGPSPRVAILREEGSNGDR
						EMADAFHLAGFEVWDVTMQDLCSGAIGLDT
1						FRGVAFVGGFSYADVLGSAKGWAAAVTFHP
1					•	RAGAELRRFRKRPDTFSLGVCNGCQLLALLG
						WVGGDPNEDAAEMGPDSQPARPGLLLRHNL
]						SGRYESRWASVRVGPGPALMLRGMEGAVLP
[VWSAHGEGYVAFSSPELQAQIEARGLAPLHW
) !						ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR
						HLAVMPHPERAVRPWQWAWRPPPFDTLTTS
						PWLQLFINARNWTLEGSC
1345	2695	Α	10396	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFQHRE
j i	<u> </u>			1		RVAMHYQMSVTLKYEIKKLIYVHLVTWLLLV
						AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI
.		1		i		LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI
'				l		YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI
				j		MYLVLVLAVQVHAWQLYYSKKLLDSWFTST
				1		OEKKHK
1346	2696	A	10398	1	718	DDFVRCGPOSAAMGASARLLRAVIMGAPGS
-	1-				-	GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT
	i		-		l	EIGVLAKAFIDOGKLIPDDVMTRLALHELKNL
	j	ļ]			TOYSWLLDGFPRTLPQAEALDRAYOIDTVINL
					l	NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK
		Į			l	TVGIDDLTGEPLIQREDDKPETVIKRLKAYED
		ĺ				QTKPVLEYYQKKGVLETFSGTETNKIWPYVY
	- (ĺ	-		1	AFLQTKVPQRSQKASVTP
			1			
1347	2697	\overline{A}	10402	153	1969	KHROENNALDMA PEILMTCOMCLIENTAICEL
1347	2697	A	10402	153	1969	KHRQENNALDMAPEIHMTGPMCLIENTNGEL VANPEALKILSAITQPVVVVAIVGLYRTGKSY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}]	j		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
İ				peptide		/=possible nucleotide deletion, \=possible
1		Ì		sequence	Į	nucleotide insertion
				1		LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
1		}				PHPKKPEHTLVLLDTEGLGDVKKGDNONDS
1	[WIFTLAVLLSSTLVYNSMGTTNQQAMDQLYY
İ						VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
ì	i					WILRDPSLDLEADGQPLTPDEYLEYSLKLTO
						GTSQKDKNFNLPRLCIRKFFPKKKCFVFDLPI
	ľ					HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
1		1				FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
1	1	}				GDLPCMENAVLALAQIENSAAVQKAIAHYD
						QOMGQKVQLPAETLQELLDLHRVSEREATEV
1	1					YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
	İ					QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK
j		ļ				PGGYCLFIOKLODLEKKYYEEPRKGIOAEEIL
	İ					QTYLKSKESVTDAILQTDQILTEKEKEIEVEC
1						VKAESAQASAKMVEEMQIKYQQMMEEKEKS
İ						YQEHVKQLTEKMERERAQLLEEQEKTLTSKL
1	ì	1	ľ			QEQARVLKERCQGESTQLQNEIQKLQKTLKK
1						KTKRYMSHKLKI
1348	2698	Ā	10404	5	892	TOLPAPLSGVLSRLOLGSGAPLLTWVOETAG
ì		l ^	10104	-		
		î	10404			VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS
		A	10404			VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
		n e	10404			VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
		C	10404			VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM
		c	10101			VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
		Α	10101			VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
		Α				VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
		Α				VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
1349	2699	A	10409	59	1184	VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL
1349	,				1	VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAILWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK
1349	,				1	VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ
1349	,				1	VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE
1349	,				1	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
1349	,				1	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
	2699	A	10409	59	1184	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM
	2699	A	10409	59	1184	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
	2699	A	10409	59	1184	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
	2699	A	10409	59	1184	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.



Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20